(Re)surveying Mediterranean Rural Landscapes: GIS and Legacy Survey Data

Robert Witcher

Department of Archaeology, Durham University, South Road, Durham DH1 3LE, UK

r.e.witcher@durham.ac.uk

Table of Contents

Summary
1 Introduction
2 Background
    2.1 Legacy survey data
    2.2 Regional survey and GIS
    2.3 Comparative survey and metadata
    2.4 GIS and comparative survey
    2.5 Social action and reflexive archaeology
    2.6 Summary
3 Case Studies
    3.1 Biferno valley – data structures
    3.2 Biferno valley – re-evaluating research questions
    3.3 Biferno valley – understanding ‘presence’
    3.4 Biferno valley – memory and movement
    3.5 Northern Lazio – chronological variability
4 Discussion
5 Conclusions
    5.1 What has the study shown?
    5.2 What are the implications for GIS research?
    5.3 What next?
    Acknowledgements
    Bibliography

Abstract

Legacy data have always been important for Mediterranean archaeologists. Over the past decade, one specific category of legacy data, that deriving from regional survey, has become particularly important. Not only has the scale of research questions become larger (requiring greater reliance on others’ data), but the surface archaeological record is deteriorating (diminishing the ability to recover good data). The legacy data from many individual surveys have now been subject to digitisation and GIS analysis, successfully redeploying data collected for one purpose within new theoretical and interpretive frameworks. However, a key research focus is now comparative survey – using the results of many different Mediterranean surveys side-by-side to identify regional variability in settlement organisation, economy and demography. In order to overcome the significant methodological differences between these surveys, attention has focused on the documentation of metadata. Yet, many legacy data lack vital information about their creation and hence how they might be (re)interpreted and compared. GIS has been advanced as an environment in which to contain, order and analyse the data necessary for comparative survey. However, there is a danger that the technology will facilitate inappropriate use of these datasets in a way that fails to acknowledge and understand the very real differences between them. Here, emphasis is placed upon the use of GIS as a space for exploratory data analysis: a process that encompasses and emphasises the integral processes of digitisation, visualisation and simple analysis for the characterisation of datasets in order to derive an alternative form of metadata. Particular emphasis is placed upon the interaction of past human behaviour (e.g. in the Roman period) and archaeological recovery (i.e. the behaviour of archaeologists in the present, or recent past); these two sets of 'social action' combine to create distinctive archaeological datasets. This search for 'contextual' metadata within individual legacy survey datasets supplements more 'formal' metadata and facilitates both interpretation of individual surveys and comparison between them. Case studies from Italy are used to illustrate the kind of iterative and incremental approach to legacy survey data advocated.

This article will appeal to: archaeologists using GIS to interpret and compare multiple legacy datasets at a landscape scale, especially Mediterranean field survey data.

Keywords: GIS, fieldwalking, comparative regional survey, metadata, settlement, Italy

1. Introduction

Over the past fifty years, hundreds of regional field surveys across the Mediterranean and Near East have documented tens of thousands of surface scatters and revealed the scale and complexity of ancient rural landscapes. The number of new and ongoing surveys attests the continuing popularity of this technique (Cherry 2003). However, the long-term erosion and deterioration of surface archaeological remains means that the results of earlier fieldwork projects are becoming more, not less, valuable over time. New methods to re-analyse and extract more meaning from these older survey data are therefore urgently required. At the same time, much attention is now focused on comparative survey, exploring inter-regional variability, particularly in areas of enhanced connectivity such as the Mediterranean. Legacy survey data are vital for this work, as no individual survey can provide the necessary coverage.

Over the last ten years, new field surveys have routinely used GIS for the storage, management, analysis and presentation of results (Gillings 2001). In the future, the importance of this technology will increase further, as mobile GIS applications take analysis into the field, and online archives and GIS allow data to be more freely exchanged. The centrality of GIS for current and future regional survey work is therefore assured. But is there a role for GIS in the study of the legacy data produced by surveys between the 1950s (and even earlier) and the 1990s?

This article provides background to the use of legacy survey data within GIS, consideration of some theoretical issues involved and case studies of practical applications concerning digitisation and analysis. In particular, it explores methodological issues of data comparability and interpretive issues concerning past social action.

Integrating the results of different field surveys is always problematical because methodologies vary significantly. This is particularly the case with legacy data as survey methodologies have evolved rapidly. The aim of this article is to demonstrate that GIS is a critical tool for the integration and re-evaluation of this massive ‘back catalogue’ precisely because it can foreground and accommodate these complexities. Armed with more nuanced appreciation of these data, GIS then provides a powerful environment within which to realise their potential in the service of new research agendas. Case studies from northern Lazio and the Biferno valley in Molise, Italy are used to illustrate specific issues.

2. Background

2.1 Legacy survey data

Archaeological surveys have been conducted around the Mediterranean for more than fifty years and ‘topographical’ surveys have an even longer history (Cambi and Terrenato 1994, 21-30). This great time-depth of antiquarian and archaeological investigation means that the Mediterranean provides an enormous archive of legacy data. By comparison with the sophisticated methodologies implemented today, many of these surveys are methodologically naïve. Issues such as site definition were frequently unconsidered and rarely documented (for an example of some of the problems, see Witcher and Craven, in press). As a result, many legacy data comprise little more than symbols on a map, perhaps associated with broad dates of occupation and/or a list of the most interesting finds. It should, of course, be emphasised that it would be anachronistic to judge such surveys by today’s standards. Many were pioneering projects, responding to new opportunities with innovative techniques. Subsequent critique and methodological developments have been fundamental, but these should not be used to condemn the methods of earlier surveyors.

The legacy of Mediterranean field survey therefore comprises large quantities of relatively low-quality data that do not provide the kind of detail and methodological rigour which many recent approaches demand. For example, without full quantification, it is difficult to assess variation in the consumption of material culture across spatial and temporal scales (Witcher 2006, 51). In many cases, it seems impossible to do more than visualise a few basic characteristics. So why persist when legacy data are badly compromised and when sophisticated new surveys can produce better data?

The answer lies partly in the changing nature of the surface archaeological record. Scatters of archaeological material in fields are usually ploughed-out of buried archaeological stratigraphy (though material may also derive from manuring; Bintliff and Snodgrass 1985). In many cases, excavation demonstrates extensive or even complete destruction of associated buried deposits. The surface scatters may therefore be the only surviving record of past activity.

However, the long-term survival of the surface archaeological record is threatened. Agricultural processes gradually abrade sherds of pottery and expose fragile materials to the elements; scatters as a whole may be ‘smeared’ and washed down hillsides. The combined effect is to reduce both the ability to recognise scatters and to recover any useful material from them. Other threats include urbanisation, road building and quarrying. Although survey has often been presented as a repeatable exercise (in contrast with excavation), in reality it is akin to ‘rescue archaeology’; indeed, John Ward-Perkins (1961, 1) explicitly initiated the South Etruria survey in response to the destruction of sites by agricultural intensification. In a gloomy assessment, Cherry (2003, 157) notes that the surface archaeological record will have disappeared altogether by *c*. 2050.

A slightly different problem is created by agricultural abatement. In the northern Mediterranean, declining subsidies and new environmental priorities mean that much land that was taken into production during the post-war years, is now reverting back to pasture or scrub and eventually woodland. As a result, there is a decline in the amount of land under cultivation and therefore less land of good surface visibility available for field survey.

Combined, the long-term degradation of the surface record and the decline of arable cultivation mean that legacy survey data form a unique resource that is growing, not diminishing, in importance. In many cases, scatters have permanently disappeared and are preserved solely ‘by record’. Other scatters have been significantly degraded since their original recognition and though resurvey may employ more systematic methodologies, the quality of material available for collection is invariably much reduced (e.g. Di Giuseppe et al. 2002).

A new generation of archaeologists has therefore begun to recognise the value of old data. In particular, many of these datasets have never been subject to any detailed analysis. The *Forma Italiae* surveys in Italy are a good example. This ongoing series of surveys aims to provide a record of the archaeological landscapes of Italy; each volume covers a single 1:25 000 *c*.10x10 km mapsheet (Cambi and Terrenato 1994, 27-32). The primary concern of the series is to catalogue and map the distribution of sites for Cultural Resource Management purposes; earlier volumes made little attempt to analyse these results (e.g. statistical tests of distribution, inter-visibility studies, etc.; recent volumes have adopted more rigorous field methodologies and analytical techniques). The unexploited potential of these earlier surveys invites renewed interest.

New research questions have also encouraged interest in old data. For example, ongoing debate about the demography of Roman Italy has found the older regional site-based surveys (e.g. South Etruria, Albegna valley) more suitable for population reconstruction than the recent small-scale, ‘off-site’ surveys (e.g. Rieti Basin) (Fentress forthcoming).

Hence, the accelerating degradation of the surface record, the limited analysis conducted by many original surveys and the evolution of archaeological ideas have all led to a growing trend in ‘second generation analysis’ of legacy datasets (for this term, see Diacopoulos 2004).

2.2 Regional survey and GIS

Regional survey and GIS are seemingly natural bedfellows (see Gillings 2001). The distribution map has long been the standard means of presenting the results of fieldwalking; from here, it has been an easy step to the enhanced flexibility and analytical power of GIS. The bulk of all early archaeological applications of GIS concerned the analysis of regional settlement patterns (e.g. Allen et al. 1990; Lock and Stančič 1995). The adoption of GIS in the context of Mediterranean regional survey has been particularly rapid. This is well-illustrated by the complete absence of any mention of GIS at the 1988 ‘Roman Landscapes’ conference (published as Barker and Lloyd 1991), soon followed by the dedication of a whole POPULUS colloquium to the topic in 1995 (published as Gillings et al. 1999).

Today, nearly all regional surveys routinely use GIS for the collection, integration, interrogation and display of results. Increasingly, survey data are digitally captured in the field with Global Positioning Systems and Personal Digital Assistants and are directly integrated with cartographical and existing archaeological data for instant analysis by mobile GIS during the course of fieldwork. However, from the beginning, legacy survey data have played a central role in GIS applications and they continue to do so today. Initially, desk-based applications were concerned with simple overlay functions (e.g. the distribution of sites in relation to soils); however, analysis has become increasingly sophisticated. For example, modelling studies and visibility analysis have been used to explore more contemporary theoretical concerns such as perception and embodiment.

A few examples of the use of legacy data within GIS stress the breadth of approach. The importance of movement for the social construction of space is explored through analysis of roads and paths (e.g. Bell et al. 2002) and the comparison of cost surfaces and viewsheds has been used to emphasise the contrasting physical and visual accessibility of rock-cut tombs (Belcher et al. 1999). Other GIS applications have considered ‘ideational’ landscapes. For example, van Hove (2004) uses GIS to consider landscape perception via Ingold’s (2000) concept of ‘taskscape’. Meanwhile, more ‘traditional’ themes such as agriculture and demography are approached in more nuanced ways. For example, Goodchild (2005) uses multi-criteria models to assess the Roman agricultural treatises; variables such as sowing ratios, labour input and yields are combined in order to assess key issues of food supply, demography and town-hinterland relations.

Finally, GIS plays a central role in the visualisation and presentation of results and data dissemination. Many regional survey projects now publish via the Internet (e.g. Davis et al. 2005) and/or deposit data with national digital archives (e.g. Given et al. 2007). The ability to share and query spatial datasets makes projects’ results dynamic – surveyors’ statements can be re-evaluated and new questions investigated. Online mapping tools such as GoogleEarth have made the manipulation of datasets even easier.

The importance of GIS for both legacy data and new regional surveys is unquestionable. As technologies and data become more freely available and easy to integrate, GIS will inevitably play an even greater role in the future. However, the vast majority of legacy survey data exist as paper catalogues. To realise the latent potential of these datasets and to subject them to new and innovative analysis requires digitisation. Unlike new survey projects where GIS design has become integral to the recording and analysis of data, the digitisation of legacy data requires a more flexible approach to accommodate problems such as inconsistent terminology or the variable accuracy and/or precision of geographical coordinates. Appropriate mechanisms to accommodate these problems must be found. Usually these solutions are a compromise between preserving the complexity of the original data and introducing sufficient order so that the data can be efficiently manipulated. Certainly the investment of time and effort necessary in order to digitise legacy data should not be underestimated; indeed, the preparation of data comprises a significant percentage of the duration of any GIS project, leading some to question the value of GIS all together (e.g. Sharon et al. 2004). However, as will be argued below, this process of data preparation needs to be reconceived.

2.3 Comparative survey and metadata

So far, discussion has focused on the analysis of legacy data from *individual* surveys. However, there is growing interest in comparing between different surveys and regions (Alcock and Cherry 2004). The potential value of comparison around and beyond the Mediterranean has long been recognised (e.g. Cherry 1983, 406), though until the last five years, progress has been slow. The most obvious stumbling block has been the diversity of survey methodologies. Simple questions such as ‘what is a site?’ have generated competing conceptual models and diverse field methods. The result has been much discussion and relatively little progress towards integrating and comparing surveys. Over the past decade, however, the achievement of this goal has been advanced by:

1. growing awareness of the variability of regional settlement through the ongoing publication of projects, including final reports by some of the larger and most ambitious surveys of the 1970s, 1980s and early 1990s (e.g. Libyan Valleys survey, Barker 1996);

2. increasing emphasis on the connectivity and interdependence of Mediterranean societies (e.g. Horden and Purcell 2000), itself partially based on recognition of the results of field survey as evidence for high levels of exchange and communication;

3. expanding availability of integrative spatial technologies (GIS, GPS, etc.) and digital data.

Comparison of regional settlement patterns around the Mediterranean during the Archaic, Hellenistic, Roman and Medieval periods should offer great insights into variable political and economic connectivity and large-scale demographic processes. Legacy data assume an important role in any such programme. No individual project can encompass the breadth of an individual country, let alone the Mediterranean and beyond, thus comparative surveyors will always need to work with legacy data.

As already discussed, these legacy data are important because the surface archaeological record is deteriorating. But they also have another particular significance for comparative survey. One of the key methodological trends over the past 25 years has been a significant increase in the intensity of coverage and a dramatic decline in the size of regions covered. The merits of this development are hotly debated (e.g. Fentress 2000; Terrenato 2004). Specifically, it is argued that this reduction in the scale, and the incompatibility of off-site methodologies, has derailed the comparative survey agenda (Blanton 2001). Some have therefore turned to the older surveys as the basis for inter-regional comparison (Blanton 2004; Fentress forthcoming); others have argued for more variety, using the older surveys as a wider framework for the greater detail provided by smaller surveys (Witcher 2006, 61-2). Notably, few have dared to reject older regional surveys outright, as this is tantamount to rejecting significant progress with comparative survey altogether (but see Ammerman 2004, 182). However this debate is resolved, it is clear that legacy data lie at the heart of comparative survey. Legacy data have to be combined with new datasets to explore large-scale questions of settlement and economy over time and significant issues of data compatibility must therefore be confronted.

How then are these methodological differences to be addressed in order that surveys can be compared? The answer is metadata; in other words, not the results of surveys, but data about the results (Wise and Miller 1997). A simple example illustrates the point: a survey using walker spacing of 10m will recover all sites more than 10m in diameter, while surveys using wider spacing will recover only a percentage; if these small scatters represented for example, burial sites, spacing of more than 10m will systematically underestimate the scale of funerary activity. Metadata therefore allow the significance of individual survey results to be understood and, in turn, allow meaningful comparison with other surveys.

The range of metadata which can be documented is varied. Some aspects of a survey can be directly measured (e.g. walker spacing, artefact sampling). These metadata can be used to analyse and compare results formally. ‘Proxy’ metadata are also common (e.g. person days per square kilometre). Although these can be measured, their relationship with survey results is not direct but can help to explain survey results (e.g. a survey of one person day per square kilometre will probably find fewer sites than a survey of twenty person days per square kilometre). In addition, to these ‘formal’ metadata, there are other forms of ‘data about data’ which defy measurement – even categorisation – but which can assist understanding. For example, survey aims have a strong influence on results through methodology and interpretation, but this influence is difficult to measure directly. Similarly, knowledge of local ceramic sequences is incredibly variable and profoundly affects the ability to date material and, hence, sites. The scale and implications of such influences are difficult to measure, but can be highly significant. In what follows, I will maintain a broad distinction between formal metadata and contextual metadata.

2.4 GIS and comparative survey

GIS was quickly identified as the ideal medium through which to facilitate comparative survey – a uniform and unifying environment in which to integrate, map and analyse data (Allen et al. 1990). However, no-one has yet created a pan-Mediterranean survey GIS. The scale of the task and the methodological issues outlined above arguably present too great a challenge to be achieved in a single step. An initial and more manageable step towards this goal is the collation of survey metadata.

The Mediterranean Archaeology GIS (MAGIS) is an ambitious initiative to collate (formal) metadata about Mediterranean surveys with the aim of disseminating information about individual surveys in order to facilitate comparisons (Foss and Schindler 2007). This project provides a welcome and constructive GIS application with which to advance the comparative survey agenda. By documenting individual surveys using a standard range of metadata, it is possible to begin to understand how they can be compared.

Meanwhile a number of projects that compare and integrate the actual datasets created by different surveys are also underway. These include the Tiber valley project (Kay and Witcher 2005; Patterson et al. 2004) and the Regional Pathways to Complexity project (Attema et al. 1998). In both of these examples, GIS plays a central role in the collation and analysis of legacy data. However, as many of the original surveys did not explicitly document or publish formal metadata, how can their results be rigorously compared?

As stressed above, we need to build up contextual metadata through a process of data characterisation; that is, to engage in 'source criticism' (see Alcock 1993, 49-53). GIS is an environment in which this understanding can be created through the processes of systematically modelling, visualising and analysing the structure of legacy data. The following case studies demonstrate this approach.

However, before considering these examples, it should be stressed that the digitisation of legacy data does not automatically make surveys comparable. When modelling individual survey results, their idiosyncratic nature can be accommodated with a degree of flexibility. Projects that integrate the results of several surveys demand greater uniformity in order that data can be systematically mapped and analysed (see Gillings 2001, 109). For example, divergent terminologies for specific types of pottery require standardisation. However, this process of lexical harmonisation does not necessarily make the data more comparable; a survey that did not collect a particular type of pottery is no more comparable with a survey which did collect this type of pottery as a result of the standardisation of terminology. Worse still, this methodological difference may be disguised by the process of standardisation (see Miller and Richards 1995); a difference in methodology may be transformed into an apparent difference in past human activity.

In using GIS for comparative survey, there is a danger that survey results are reduced to the lowest common attributes (where, when, what); this may encourage simplistic and inappropriate comparison by disguising real differences in data structures (how have spatial location, date and site interpretation been achieved?). As an initial step, GIS applications should explore the complexity of datasets rather than reducing them to (apparently) comparable basic attributes. For example, we might develop innovative ways of visualising and accommodating variation in data quality (e.g. in spatial coordinates; Boldrini 2007). Such work should be understood to be concerned with data characterisation for the identification of contextual metadata.

In summary, increasing interest in comparative survey has provided new impetus for the use of legacy survey data. Variability of methodologies raises significant issues of data compatibility. GIS is not a simple panacea to these problems; indeed, arguably it could aggravate the situation by masking incompatibilities. However, GIS can play a positive role: the processes of digitisation, visualisation and spatial analysis all assist in the creation of contextual metadata. In turn, this facilitates more informed interpretation and comparison.

2.5 Social action and reflexive archaeology

The process of source criticism is essentially a reflexive archaeological practice (Hodder 1999, 81-104). Reflexive archaeology encourages greater awareness of the role of the archaeologist in the creation of archaeological data. Whether surveying landscapes or digging trenches, the actions of archaeologists shape their datasets; for example, decisions on how to sample artefact scatters or how to define a 'site'. This does not mean there is no reality 'out there', but simply that there is no objective way of recording it. Hodder and others have called for greater reflexivity about the nature of archaeological data and, in particular, the illusory distinction between objective description (fieldwork, cataloguing) and subjective interpretation. This is, of course, of particular relevance when using legacy data as initial data capture and use may be separated from secondary re-use by decades, and filtered through various media such as archives and databases.

To date, most GIS applications have focused on studying the behaviour of people in the past, whether through settlement preferences, site intervisibility or movement. However, GIS also offers an environment in which to study the archaeological process in the present (or, in the case of legacy data, the very recent past); in other words, to develop a reflexive archaeology. To date, such analysis has focused on formal metadata (e.g. surface visibility), but a range of other contextual metadata can be addressed. For example, Rajala et al. (1999) have explored the influence of developing aims and methods on the types and distribution of sites in the *ager Faliscus*, north of Rome. Such work stresses the importance of the recognition that the actions of people in the past and the actions of archaeologists in the present are not separate areas of enquiry. The deposition and recovery of archaeological materials interact to create the specific character of archaeological datasets. For example,

- A survey's reliance on finewares to identify and date site activity will mean that a period which uses few or no finewares will be 'invisible'. This may, for example, partly explain the scale of decline in settlement activity during late Antiquity in the western Mediterranean;
- A decision to use decorative stone or other architectural features to define high-status settlement sites will prioritise the recognition of groups and individuals who invest in fixed rather than portable material culture.

The study of this relationship is particularly important when using legacy data because the actions of past archaeologists may differ significantly from those of contemporary archaeologists. Survey methodology has evolved rapidly and the actions of archaeologists 10, 25 and 50 years ago are likely to be very different. Current surveyors can think reflexively about how their own methodologies impact their data; for most legacy data, the opportunities for such reflection have been lost (Kintigh 2006).

2.6 Summary

This article does not propose a specific suite of GIS applications for legacy data or any high-level theoretical framework. Rather it argues that current analyses are supplemented in several ways:

Firstly, more emphasis should be placed on source criticism (i.e. the creation of contextual metadata through a process of data characterisation) as a means of understanding data, including particular attention to the interaction of past and present action (i.e. the behaviour of people in the past and of the archaeologists studying them).

Secondly, more value should be placed on the description and visualisation of data within GIS as a means to identify contextual metadata. The charge of 'pretty maps' is powerful in an age of interpretive archaeology (see Fisher 1999, 8). However, the importance of exploring data structures and defining characteristics should not be underestimated when dealing with legacy data. It is easy to imagine that survey data are objective 'facts', especially when available digitally 'off the shelf'; GIS should help to stress the diversity of data structures, not gloss over incompatibilities.

Thirdly, the process of data collation, preparation and entry should be rethought as an integral part of data characterisation. Digitisation is a valuable part of analysis and not just the frustrating process of 'getting the data in' or an inconvenient formality before the real analytical work starts. It is also an iterative and incremental process. It is worth stressing this point for two reasons: firstly, the importance of this phase is rarely discussed (e.g. Chapman's (2006) *Landscape Archaeology and GIS* addresses data sources, but provides no explicit discussion of data entry as an integral part of the analytical process). Second, the growing (and welcome) availability of legacy data in digital formats means that this process is increasingly completed by other people; it is therefore ever easier to underestimate the range and significance of decisions and assumptions involved not only during the original fieldwork, but also during the process of digitisation.

In summary, general approaches to GIS have begun to engage with theories of social action in the past. Meanwhile, the growing reflexivity of archaeology as a whole is encouraging greater attention towards the role of the archaeologist in shaping archaeological datasets. Legacy data require particular attention because the theories and methods of survey archaeology have rapidly evolved. GIS offers an environment in which the spatial structure of the archaeological record can be assessed – in terms of the social actions of people in the past and the archaeologists who study them. The application of GIS to these issues does not require a revolution in current approaches, but it does require a change in the way digitisation, comparison and visualisation of data are conceived.

3. Case Studies

3.1 Biferno valley – data structures

The first case study concerns the Biferno valley in the Molise region of central Italy. The results of two surveys will be discussed: the Biferno valley survey (Barker 1995a; 1995b) and the *Forma Italiae* survey of the area around the Samnite/Roman town of Larinum (modern Larino) (De Felice 1994) (Fig. 1). The former sampled *c*. 400km^2 across the entire river valley (*c*. 20%); the latter covered a *c*. 100km^2 block at 100%. The two surveys were independent; the fieldwork for the *Forma Italiae* was conducted intermittently between 1969 and 1989 (De Felice 1994, 11); the Biferno valley survey was completed between 1974 and 1978. The final publications appeared in consecutive years; neither project makes reference to the other. By contemporary standards, neither survey presents detailed metadata. Of particular interest is an area of spatial overlap between the two surveys which offers an opportunity to examine two independent archaeological surveys of the same landscape side-by-side.

Both surveys published detailed gazetteers in paper form; while it is straightforward to access data about individual sites, it is difficult to appreciate broader patterns and to evaluate the surveyors' interpretations. To this end, the data from both surveys were digitised. In addition, a range of environmental data was collected including a DEM, hydrology and soils.

The surveys present their data in distinct formats. The Biferno valley survey uses a systematic format with information divided into 'byte-sized' pieces making it relatively straightforward to normalise and model within a relational database (Table 1). The free-text format of the *Forma Italiae* gazetteer requires a more flexible approach that is sensitive enough to contain the maximum information without compromising the resulting database's analytical potential (Table 2).

The data were therefore entered into separate databases, but wherever possible using the same structures and terminologies (both surveys, for example, used the 'site' or 'findspot' as the basic unit of record). The process of data entry involved rigorous checks and the enforcement of referential integrity (e.g. material had to relate to a known site); this identified a number of errors which were not apparent in the paper gazetteers.

Once the attributes of these sites had been entered into databases, their grid references were derived in order that the two surveys could be intersected spatially and 'duplicate' records identified. Figure 2 illustrates the area of spatial overlap between the surveys. In order to identify which *Forma Italiae* findspots lay in close proximity to Biferno valley survey findspots, the latter were buffered (i.e. an area defined with a 150m radius around each findspot). These buffer areas were then overlaid onto the *Forma Italiae* findspots and those which fell within these 150m buffers were isolated for further consideration. This process identified 20 (out of 79, i.e. 23.5%) Biferno valley survey records with one or more *Forma Italiae* records within 150 metres; in total, 31 (out of 128, i.e. 19.5%) *Forma Italiae* records (excluding urban material) were found to lie within 150 metres of a Biferno valley survey findspot.

The attributes of each potential duplicate pair or group were then compared. In those cases where records clearly referred to the same archaeological feature, cross-references were created. The remaining examples were visualised in relation to slope (derived from the DEM) in order to assist interpretation. For example, thin, low-density ('sporadic') scatters found directly below large sites on steep slopes were considered likely to represent either an extension of the site or erosion of it. Cross-references were therefore created between these archaeological features. Where doubt remained over the relationship between records, they were kept separate. As a result of this process, it appears that the two surveys document the surface archaeology in different ways. The *Forma Italiae* survey uses approximately 50% more records to define the same features as the Biferno valley survey. In other words, on average, what the Biferno valley survey would record as 10 findspots, the *Forma Italiae* survey would record as 15. Recognition of this process of 'lumping' or 'splitting' is clearly significant if surveyors wish to undertake quantitative comparisons of site numbers and densities between surveys and regions.

Within the area walked by both surveys, how can we account for the fact that less than 25% of findspots were the same? If we assume that the archaeological record remained unchanged (a significant assumption, see below), then the differences are to be explained by the diverse aims and methods of the two surveys and the field conditions they experienced. The *Forma Italiae* represents a comprehensive catalogue of all archaeological evidence for Cultural Resource Management purposes compiled over many years; in contrast, the Biferno valley survey adopted a larger regional focus and aimed for an internally consistent sample through a briefer series of field seasons. *A priori*, it might be assumed that the *Forma Italiae* data would create a more comprehensive 'baseline' which the Biferno valley survey was only able to sample as part of its wider regional approach. However, this is not the case; there are 97 findspots not recorded by the Biferno valley survey and 59 findspots *not* recorded by the *Forma Italiae*. In other words, two different surveys, working independently, were able to identify a common group of findspots (recorded as 20 findspots by the Biferno valley survey, and as 31 findspots by the *Forma Italiae*), plus larger numbers unidentified by the other survey. One likely explanation is variable accessibility to fields and changing surface visibility within them. Despite the spatial overlap of the surveys, the surveyors may have encountered very different conditions within and between years (Ammerman 1985).

Another consideration is that the nature of the archaeological record itself may have changed. For example, the rapid erosion of surface material was noted at several sites identified and then revisited by the Biferno valley survey between 1974 and 1978 (Lloyd and Barker 1981, 290); it is possible that the nature of surface archaeology changed even more considerably over the extended duration of the *Forma Italiae* survey. This issue also raises a broader question of whether a definitive map should be the goal of (comparative) survey. Two surveys of the same region are likely to record different results. How then is it possible to make a definitive map, marking each site as a dot? A longer-term goal must be to use GIS to develop new ways of representing ancient landscapes at a variety of scales; not a single, definitive atlas, but a range of alternatives (Mattingly and Witcher 2004).

I have discussed the detail of data preparation at length in order to emphasise that this stage is not simply a preliminary step but integral to data analysis. Such detailed data handling facilitates important impressionistic understanding about the character of datasets. In other words, this is part of the process of source criticism. The generation of contextual metadata through an appreciation of data characteristics can subsequently inform more formal analysis. For example, through modelling and inputting data, it became clear that the Biferno valley survey contained its results within a limited number of interpretive categories; in contrast, the *Forma Italiae* was more flexible, though often reluctant to apply any interpretation at all. Similarly, it was noted that the *Forma Italiae* recorded more findspots overall, but this should be balanced against its tendency to use more records to document the same phenomena as the Biferno valley survey. Hence, data preparation can generate insight into the characteristics of datasets. Such work should be seen as an inherent part of a GIS approach to legacy survey data.

3.2 Biferno valley – re-evaluating research questions

Once the data have been digitised, further visualisation and more formal analysis can proceed. A useful approach is to re-evaluate original survey aims. For example, the Biferno valley survey was conducted during the 1970s and published during the 1990s (Barker 1995a; 1995b); hence, as with many large projects, significant theoretical and methodological advances had been made between the survey’s initiation and its publication. The surveyors’ original research questions concerned processual interests such as the relationship between settlement and environmental resources. The final report reconceived these research questions within an *Annales* framework placing more emphasis on human agency, particularly for historical periods (see Barker 1991).

It has been suggested that the time lapse between large surveys commenced in the 1970s and their analysis and publication today can lead to a situation in which GIS is used in the uncritical service of outdated theoretical perspectives (Gillings 2001, 111). However, GIS offers a tool not simply to realise or repeat original research questions, but also to re-evaluate them. Research aims have significant influence on the data collected through the methodologies used. Consideration of original survey aims enhances understanding of the character of results. Such analysis should not be the sole aim of a GIS approach to legacy data, but can form an informative part of it.

The Biferno valley survey was particularly concerned with the relationship between people and environment. An interim discussion (Barker et al. 1978, 45) argued that arable cultivation was limited to the lighter, free-draining sandy soils of the ridges around the town of Larinum and on the edge of the river plain (i.e. mixed sands/clays, limestone). The heavier, silt soils of the plain itself (i.e. alluvium) were considered to be more difficult to cultivate and, as in more recent times, were devoted to pastoral exploitation. Consequently, sites were located on the higher, lighter soils; artefact scatters on the plain were argued to represent annual transhumance camps (Barker et al. 1978, 48). Re-assessment of the distribution of findspots in relation to soils would be an extremely time-consuming exercise using the paper gazetteer; moreover, the report provides no quantification of the extent of each soil type and therefore the *significance* of any patterning could not be assessed. Such analysis is a simple matter to address via GIS.

Within the case study area, the outline of each survey was overlaid on the soil map to establish the extent of each soil class sampled (Table 3). The number of findspots located on each soil type was then derived (Fig. 3). A one-sample chi-squared test was used to assess whether the number of observed findspots per soil type was significantly higher than expected. The results confirm a significant relationship (Table 4); there are almost double the number of expected findspots on the difficult-to-work (Pleistocene) alluvial terraces. An association between the number of sites and soil type is not incompatible with the original interpretation, but the unexpected concentration of findspots on the supposedly difficult-to-use soils gives a different emphasis; in simple quantitative terms, this was a favoured area. This association can be combined with other observations. Many of these findspots are extensive, low density scatters usually of a single period, with evidence for pottery but not tile. Barker *et al*. (1978) argue that these represent pastoralists' camps, but would these have been so widely-dispersed if this were a wooded and poorly drained (even malarial) landscape? An alternative interpretation might be that these scatters derive from manuring as part of the long-term cultivation of the plain (Bintliff and Snodgrass 1985); the percentages of (datable) findspots which make use of this soil type are broadly stable from the Iron Age through to the Roman period (Fig. 4) suggesting that the exploitation of this resource waxed and waned with broader shifts in settlement numbers, rather than the introduction of new technologies or subsistence strategies to overcome environmentally-determined constraints. Assessment of the comparative abrasion of material from these findspots might prove informative; sherds derived from manuring might be significantly more abraded than sherds derived from stratigraphic contexts (in the Biferno valley, many sites were fieldwalked soon after the introduction of deep ploughing in the region, reducing the abrasion of material recently derived from stratigraphic contexts).

Simple GIS overlay functions allow a quick evaluation of the survey's original aims, results and interpretations. In this case, it is possible to corroborate a relationship between soils and settlement, but the distribution of settlement suggests a different emphasis. Interpretation might be revised to focus less on the environmental constraints on agricultural exploitation. For example, the location of villas might be determined not by the lighter soils above the plain, but by the elevated topography which could provide appropriate panoramic views for elite rural retreats. The relationship between soils, topography and viewsheds might be formally investigated via further GIS work.

3.3 Biferno valley – understanding 'presence'

A frequent problem when using legacy survey data is the lack of any record of recovery conditions (e.g. which areas were walked and which were not, variable visibility, etc.). As a result, we are not dealing with the presence and absence of sites or material culture; we are dealing with 'presence' alone. This has significant implications. A distribution map of settlement in any particular period is next to useless if we have no understanding of whether gaps or clusters are meaningful. Simple visualisation techniques can help to improve understanding.

By plotting particular subsets against the backdrop of all findspots, it is possible to gain some appreciation of whether patterning is meaningful. For example, neither visibility nor the extent of areas walked were recorded by the Biferno valley survey; it is therefore impossible to identify significant patterns such as high visibility areas which produced no surface material. Figure 5 shows the distribution of Roman villas across the Biferno valley visualised against a backdrop of all other findspots. This suggests, for example, that fieldwork was able to locate plenty of findspots in the upper valley, so the absence of villas in this area may be genuine. A simple distribution map of villas, or villas and contemporary settlement, or even villas and sampling transects, would not necessarily have permitted us to come to this conclusion. Presence data are difficult to use and require multiple visualisations to appreciate their significance.

Another approach to the 'presence-only' problem is to map findspots of one particular period in relation to findspots of preceding or subsequent periods. By plotting findspots in terms of continuity, abandonment and foundation, aspects of spatial patterning are better understood because there is some appreciation of 'absence'. For example, within the Larinum case study area, the transition between the Iron Age and Samnite periods is marked by the continuity of settlement in the hilly areas south of Larinum, with a notable spread of new sites on the lowest hills and plain to the north of the town (Fig. 6). This visualisation can also be combined with evidence for scatter size and density to suggest a series of small, but continuing sites in the hills, and larger but low density scatters in the lower hills and on the plain. This spatial reorganisation may be associated with the emergence of Larinum as the primary regional centre on the basis of enhanced control and exploitation of the agricultural territory to its north.

Legacy survey data present particular problems of use. Whether considering the relationships of settlement location with soil types or with viewsheds, it is important to have some appreciation of the spatial significance of site distribution. GIS can map anything – no matter how incomplete or biased. It is therefore essential to appreciate how legacy data have been created in order to evaluate whether or not patterns are of any significance. Basic visualisation techniques present one often underappreciated approach.

3.4 Biferno valley – memory and movement

A key issue in the use of legacy (survey) data is the degree to which data collected for one purpose can be used in the service of different research questions. In many cases, the original aims and methods mean that the data cannot be used in this way (see Chippendale 2000). For example, most pre-1980 surveys used the 'site' as the basic unit of record; it is therefore impossible to redeploy these data in order to address off-site activities such as manuring. None the less, there are still opportunities to explore recent theoretical concerns using legacy data, for example, visibility and viewshed studies have been particularly popular (e.g. Belcher et al. 1999).

Memory has formed a focus for much contemporary research within (landscape) archaeology. Individuals and groups within societies derive cultural memories from, or attach them to, elements within the landscape in order to negotiate identity and power (Alcock and van Dyke 2003). Older surveys often overlooked or dismissed the significance of such cross-temporal connections; maps of Roman settlement, for example, rarely include pre-Roman monuments which still stood in the landscape (Witcher 2006, 55). While much work on memory has concentrated on monuments, the archaeological landscapes of the Mediterranean are dominated by scatters of pottery derived from domestic and agricultural sites. However, these sites also serve to convey the historical depth and continuity of landscapes and people's relationships with them.

Again, simple visualisation techniques, such as plotting sites by the number of periods of occupation, can reveal informative alternative geographies. For example, in the lower Biferno valley, a series of well-dispersed, long-term settlement sites are surrounded by clusters of short-lived, single-phase, sites (Fig. 7). This might suggest that the underlying social and economic structure of this landscape was more stable than the radical changes implied by settlement figures alone (Fig. 8). These long-lived sites might be considered as stores of social memory and power, actively maintained by groups and individuals, perhaps even when other more suitable locations could be found. Further, the longest-lived sites tend to be higher up the settlement hierarchy (Table 5). This need not be the case; for example, villas might have been considered to be culturally intrusive and therefore vulnerable to wider economic and political shifts. Instead, the longevity of these sites suggests that they emerged from, and made cultural reference to, an existing system rather than being imposed upon it.

Another approach to the cultural depth of landscapes is to explore issues of embodiment and experience (Witcher 1998). Across ancient Italy, a close relationship between topography and communication routes was common; frequently, tracks tended to follow ridges and watersheds (e.g. in South Etruria, Kahane et al. 1968). Around Larinum, a close relationship can also be observed between sites and elevated topographical locations (Fig. 9). How was the experience of moving through this landscape structured by the relationship between topography, communication routes and the distribution of sites? In order to journey through this landscape, the DEM was processed to show the convexity (e.g. ridges) or concavity (e.g. valleys) of the landscape. The main ridge from the plain up to Larinum was then considered in terms of the sequence of findspots along its route (Fig. 10). During the Roman period, travellers moving south from the plain up the ridge to the town would have moved through a landscape occupied by isolated rural settlements, before passing Monte Arcano, the focus of an extensive Iron Age and Samnite cemetery (Coarelli and La Regina 1993, 301; Di Niro 1991, 131). A series of low stone mounds (Barker and Suano 1995, 172) advertised the past function of the site, while its overgrown character indicated its antiquity. Continuing along the ridge, travellers passed through further areas devoted to agricultural activity, before passing another archaic necropolis of tumuli at Colle San Pietro. Not far beyond, immediately outside the town, were cemeteries of the 2nd/1st centuries BC. This area formed an extensive funerary zone throughout the late Samnite and Roman periods. Having passed the disused ancient cemeteries further down the ridge, this centralisation of contemporary funerary activity immediately outside the town's boundary (the *agger*), and its new monumental forms, were impressed upon the traveller. Old foci of veneration and political order were left physically and socially peripheral as formerly distinct activities (habitation, burial, ritual and monumentalisation) were centralised at Larinum. If the emergence of the town during the early Samnite period was associated with the dispersal of landscape activity (see Understanding Presence), the early Roman transformation of Larinum involved a concentration of social power reinforced through a re-centralisation of people and activities.

3.5 Northern Lazio – chronological variability

A final example of the use of GIS with legacy data uses the evidence from two surveys in northern Lazio: the South Etruria survey and another *Forma Italiae* volume. The South Etruria survey is one of the pioneering Mediterranean regional surveys covering *c*. 1000km^2 of territory to the north of Rome (Potter 1979); here attention will focus on just one small sub-project, the Eretum survey (Ogilvie 1965). Subsequently, many other topographical and regional field surveys have been conducted in this region, including a *c*. 100km^2 survey by Pala (1976) around ancient Nomentum (Fig. 11).

The longevity and intensity of archaeological survey in this area provides plenty of examples of the problems and potential of comparing legacy data within a GIS environment. Digitisation of these surveys' results allows settlement patterns to be queried and mapped, but it does not mean that it is possible simply to forget about the disparate methodological origins of the data. One of the key problems for comparative survey concerns the disparate chronological frameworks used. Not only might regional chronologies differ, but individual surveys might (and, in fact, often do) define their own dating schemes (e.g. 300 BC–1 BC vs. 250 BC–27 BC vs. 'Republican'). In terms of comparison, these schemes are simple to record with formal metadata (though practicable strategies for integration are less straightforward). However, there are other aspects of survey chronology that are more difficult to document. How far has a survey made (full) use of fine- and coarseware typologies? How does knowledge of (local) ceramic sequences vary across time and space? As a result, the comparability of chronological information from legacy data is far more problematical than the more familiar problem of comparing spatial information. However, GIS can help to visualise these problems and therefore to manage them.

For example, in Figure 12, findspots are plotted according to dating resolution – generic Roman (*c*.350 BC–AD 250), imperial (*c*.31 BC–AD 250) and early imperial (*c*.31 BC–AD 100). The symbols are 'stacked up'; the smaller and darker the symbol, the more closely it is possible to date the findspot. The distribution of symbols indicates clear differences in levels of chronological resolution achieved by the Eretum survey and the Nomentum (*Forma Italiae*) survey. Here, the greater dating precision of the Eretum survey is the result of a strong emphasis on ceramics for dating, while the more generic chronological identifications of the Nomentum survey reflect its reliance on structural remains and inscriptions. A simple map of early imperial findspots (with none of the more generically dated sites) would indicate a strong, but misleading, cluster of activity around the site of Eretum which might (wrongly) suggest Eretum was a vibrant market centre at this date. Such visualisation techniques are a powerful means of data characterisation. The resulting contextual metadata are critical for the analysis, interpretation and comparison of legacy data.

4. Discussion

In each of the case studies presented above, the application of GIS to legacy data has added nuance to understanding of survey results. In particular, it is argued that such applications need to be flexible and to focus upon not only past human behaviour, but also the practice of archaeology itself and the effects this has on data structure and interpretation. Comparative survey requires a wide range of formal metadata (e.g. walker spacing) with which to evaluate the results of individual surveys and to compare one survey with another.

Such metadata are scarce for most legacy surveys. However, through digitisation, visualisation and simple analysis, it is possible to enhance understanding of such data. For example, consideration has been given to how two different surveys record the same landscape and whether the distribution of sites is affected by visibility. Further, it has been possible to assess a survey's original research questions and to re-evaluate the validity of the surveyors' conclusions. It has also been possible to approach new research topics such as memory and movement even though these data were not collected with these questions in mind.

As a result, we have developed new understanding of past social behaviour in the Biferno valley. For example the emergence of Larinum as an urban centre was associated with enhanced exploitation of its immediate hinterland, in particular the territory to the north. However, its subsequent transformation into a Roman town involved a recentralisation of social activities (e.g. the thinning-out of rural settlement and refocusing of monumentality and burial at the site of Larium itself; see Memory and Movement above). More generally, the transformation of settlement organisation was incremental; rather than an unstable landscape of many short-lived sites, it is possible to identify a network of long-term settlement foci around which more temporary activities occurred. Arguably these long-lived sites formed 'stores' of memory and social capital through which change such as villa-building could be mediated. Individually none of these issues revolutionises understanding of these surveys; cumulatively such nuances refine understanding and permit broader reassessments.

5. Conclusions

5.1 What has the study shown?

This study has explored the problems and potential of using legacy survey data within a GIS environment. The emphasis has not been to promote a single analytical framework, but rather a more reflexive approach with which to supplement the existing diversity of applications.

Instead of using GIS as a convenient mechanism for simplifying and comparing data, this study has attempted to use GIS to explore the unevenness of legacy data and to generate contextual metadata as a substitute for more formal records. It has been argued that all datasets are artefacts of the methodologies used to collect them. If legacy data are to be used with any degree of sophistication, it is vital to consider how the conceptualisation and execution of fieldwork has influenced results; often these methodological decisions interact directly with the nature of the archaeological record. For example, a transect survey of a nucleated settlement pattern may achieve rather limited results. With legacy data, often questions arise which cannot be answered because of data collection methods. For example, is the lack of early imperial settlement in the Biferno valley due to rural poverty, settlement nucleation or population decline? The decision to 'grab' sample artefacts makes it difficult to assess which of these possibilities is the most plausible. However, by understanding the limitations of such data, it is possible to develop practicable field strategies for targeted resurvey to resolve such issues and bring about meaningful comparison (for an example, see Di Giuseppe et al. 2002).

5.2 What are the implications for GIS research?

Legacy survey data are not objective archives of facts and figures; they require careful and critical use. When the data from one survey are to be compared with another, these issues become even more significant. When the data come in a pre-digitised format, there is a real danger that these differences can be underestimated if unaccompanied by a wealth of formal metadata. If GIS is to become part of the solution, rather than potentially making the problem worse, then it is important to acknowledge that GIS is not a neutral tool but an affective approach (Gillings 2001, 109). It is an environment in which the data are transformed, interpreted and reinterpreted. The application of GIS should be an iterative process which focuses on enriching knowledge of data structures – with equal emphasis on the role of archaeologists and the behaviour of people in the past. Letting legacy data 'speak for themselves' risks failing to acknowledge the very real impact of the archaeologist in creating the data – whether the original surveyor or subsequent generations of archaeologists who will increasingly make use of them.

5.3 What next?

GIS is already established as an integral part of the analysis and interpretation of survey. Legacy data have been heavily used since the emergence of archaeological GIS in the early 1990s and their use continues to grow. This article argues that, alongside sophisticated modelling and analytical methods, the processes of digitisation and visualisation should be recognised as integral to the improved understanding of legacy data. Small-scale, flexible and reflexive GIS environments have much to offer the longer-term aim of large-scale GIS comparison of settlement, economy and population across the Mediterranean. An intermediate step is to explore the opportunities provided by overlapping and adjacent surveys for identification and refinement of basic techniques of data characterisation. The collation of formal metadata for legacy survey data is already underway; what is needed now is consideration of the other types of metadata that can be extracted, and further thought about how these can be documented, shared and used.

Acknowledgements

Thanks to Pim Allison for the invitation to contribute and to Alice Hiley for her assistance during the writing of this article.

Bibliography

Alcock, S.E., 1993 *Graecia Capta: The Landscapes of Roman Greece.* Cambridge.

Alcock, S. and Cherry, J. (eds), 2004 *Side-by-Side Survey. Comparative Regional Studies in the Mediterranean World*. Oxford.

Alcock, S.E. and Van Dyke, R.M. (eds), 2003 *Archaeologies of Memory*. Malden, Mass.

Allen, K.M.S., Green, S.W. and Zubrow, E.B.W. (eds), 1990 *Interpreting Space. GIS and Archaeology*. London, New York and Philadelphia.

Ammerman, A.J., 1985 'Plow-zone experiments in Calabria, Italy', *Journal of Field Archaeology* **12**, 33-40.

Ammerman, A.J., 2004 ‘Farewell to the Garden of Eden: survey archaeology after the loss of innocence’, in M. Iacovou (ed.), *Archaeological Field Survey in Cyprus: Past History, Future Potentials.* London, 177-82.

Attema, P.A.J., Burgers, G.-J., Kleibrink, M. and Yntema, D.G., 1998 'Case studies in indigenous developments in early Italian centralization and urbanization: a Dutch perspective', *Journal of European Archaeology* **1**, 326-81.

Barker, G. 1991 ‘Two Italys, one valley: an Annaliste perspective’, in J. Bintliff (ed.), *The Annales School and Archaeology.* Leicester, 34-56.

Barker, G. (ed.), 1995a *A Mediterranean Valley. Landscape Archaeology and Annales History in the Biferno Valley*. London.

Barker, G. (ed.), 1995b *The Biferno Valley Survey. The Archaeological and Geomorphological Record*. London.

Barker, G. (ed.) 1996 *Farming the Desert. The UNESCO Libyan Valleys Archaeological Survey. Volume One: Synthesis*. Paris / Tripoli / London.

Barker, G. and Lloyd, J. (eds), 1991 *Roman Landscapes: Archaeological Survey in the Mediterranean Region*. London.

Barker, G., Lloyd, J. and Webley, D., 1978 'A classical landscape in Molise', *Papers of the British School at Rome* **46**, 35-51.

Barker, G. and Suano, M., 1995 ‘Iron Age vhiefdoms, c. 1000-500bc’, in G. Barker (ed.), *A Mediterranean Valley. Landscape Archaeology and Annales History in the Biferno Valley.* London. 159-80.

Belcher, M., Harrison, A. and Stoddart, S., 1999 ‘Analysing Rome's hinterland’, in M. Gillings, D.J. Mattingly and J. Van Dalen (eds), *Geographical Information Systems and Landscape Archaeology.* Oxford, 95-102.

Bell, T., Wickham, A. and Wilson, A., 2002 'Tracking the Samnites: landscape and communications routes in the Sangro Valley, Italy', *American Journal of Archaeology* **106**, 169-86.

Bintliff, J. and Snodgrass, A., 1985 'The Cambridge/Bradford Boeotia expedition: the first four years', *Journal of Field Archaeology* **12**, 123-61.

Blanton, R.E., 2001 'Mediterranean myopia', *Antiquity* **75**, 627-29.

Blanton, R.E., 2004 ‘Settlement pattern and population change in Mesoamerican and Mediterranean civilizations: a comparative perspective’, in S.E. Alcock and J.F. Cherry (eds), *Side-by-Side Survey. Comparative Regional Studies in the Mediterranean World.* Oxford, 206-40.

Boldrini, N., 2007 ‘planning uncertainty: creating an artefact density index for North Yorkshire, England’, *Internet Archaeology* **21**. http://intarch.ac.uk/journal/issue21/boldrini_index.html. Accessed 6 September 2007.

Cambi, F. and Terrenato, N., 1994 *Introduzione all'archeologia dei paesaggi.* Rome.

Chapman, H.P., 2006 *Landscape Archaeology and GIS.* Stroud.

Cherry, J.F., 1983 ‘Frogs around the pond: perspectives on current archaeological Survey projects in the Mediterranean area’, in D.R. Keller and D.W. Rupp (eds), *Archaeological Survey in the Mediterranean Area.* Oxford, 375-416.

Cherry, J.F., 2003 ‘Archaeology beyond the site: regional survey and its future’, in J.K. Papadopoulos and R.M. Leventhal (eds), *Theory and Practice in Mediterranean Archaeology: Old World and New World Perspectives.* Los Angeles. 137-59.

Chippendale, C., 2000 'Capta and data: on the true nature of archaeological information'. *American Antiquity* **65**, 605-12.

Coarelli, F. and La Regina, A., 1993 *Abruzzo - Molise.* Bari.

Davis, J.L., Alcock, S.E., Bennet, J., Lolos, Y., Shelmerdine, C.W. and Zangger, E., 2005 *The Pylos Regional Archaeological Project: Internet Edition*. http://classics.uc.edu/prap. Accessed 6 September 2007.

De Felice, E., 1994 *Larinum.* Florence.

Diacopoulos, L., 2004 'Investigating social complexity through regional survey: “second-generation” analysis of Bronze Age data from the Canadian Palaipaphos Survey Project, Southwestern Cyprus', *Journal of Mediterranean Archaeology* **17**, 59-85.

Di Giuseppe, H., Sansoni, M., Williams, J. and Witcher, R.E., 2002 'The *Sabinensis Ager* revisited: a field survey in the Sabina Tiberina', *Papers of the British School at Rome* **70**, 99-150.

Di Niro, A., 1991 ‘La zona frentana tra IV e I secolo a.C.’, in S. Capini and A. Di Niro (eds), *Samnium. Archeologia del Molise*<.em>. Rome, 131-34.

Fentress, E., 2000 ‘What are we counting for?’, in R. Francovich and H. Patterson (eds), *Extracting Meaning from Ploughsoil Assemblages.* Oxford, 44-52.

Fentress, E., forthcoming ‘Peopling the countryside: Roman demography in the Albegna Valley and Jerba’, in A. Bowman and A. Wilson (eds), *Approaches to Quantifying the Roman Economy*. Oxford.

Fisher, P.F., 1999 ‘GIS: today and tomorrow?’, in M. Gillings, D.J. Mattingly and J. van Dalen (eds), *Geographical Information Systems and Landscape Archaeology.* Oxford, 5-12.

Foss, P. and Schindler. R. (eds), 2007 *MAGIS: Mediterranean Archaeological GIS*. DePauw University. http://cgma.depauw.edu/MAGIS. Accessed 6 September 2007.

Gillings, M., 2001 ‘The utility of the GIS approach in the collection, management, storage and analysis of surface survey data’, in J. Bintliff, M. Kuna and N. Venclova (eds), *The Future of Archaeological Field Survey in Europe.* Sheffield, 105-20.

Gillings, M., Mattingly, D.J. and van Dalen, J. (eds), 1999 *Geographical Information Systems and Landscape Archaeology*. Oxford.

Given, M., Corley, H. and Sollars, L., 2007 'Joining the dots: continuous survey, routine practice and the interpretation of a Cypriot landscape', *Internet Archaeology* **20**. http://intarch.ac.uk/journal/issue20/taesp_index.html. Accessed 6 September 2007.

Goodchild, H., 2005 ‘Modelling agricultural production. A methodology for predicting land use and populations’, in J. Haldon (ed.), *General Issues in the Study of Medieval Logistics: Sources, Problems and Methodologies.* Leiden, 199-228.

Hodder, I., 1999 *The Archaeological Process.* Oxford.

Horden, P. and Purcell, N., 2000 *The Corrupting Sea. A Study of Mediterranean History.* Oxford.

Ingold, T., 2000 *The Perception of the Environment. Essays on Livelihood, Dwelling and Skill.* London and New York.

Kahane, A., Murray Threipland, L. and Ward-Perkins, J., 1968 ‘The *Ager Veientanus*, north and east of Veii’, *Papers of the British School at Rome* **36**, 1-218.

Kay, S. and Witcher, R.E., 2005 'The Tiber Valley Project. The role of GIS and databases in field survey data integration and analysis', *Archeologia e Calcolatori* **16**, 113-37.

Kintigh, K.W., 2006 'The promise and challenge of archaeological data integration', *American Antiquity* **71**, 567-78.

Lloyd, J. and Barker, G., 1981 ‘Rural settlement in Roman Molise: problems of archaeological survey’, in G. Barker and R. Hodges (eds), *Papers in Italian Archaeology 2.* Oxford, 289-304.

Lock, G. and Stančič, Z. (eds), 1995 *Archaeology and Geographical Information Systems: a European Perspective*. London.

Mattingly, D.J. and Witcher, R.E., 2004 ‘Mapping the Roman world: the contribution of field survey data’, in S.E. Alcock and J. Cherry (eds), *Side-by-Side Survey. Comparative Regional Studies in the Mediterranean World.* Oxford, 173-86.

Miller, P. and Richards, J., 1995 ‘The good, the bad, and the downright misleading: archaeological adoption of computer visualisation’, in J. Huggett and N. Ryan (eds), *Computer Applications and Quantitative Methods in Archaeology - CAA94.* Oxford, 19-22.

Ogilvie, R.M. 1965. Eretum. *Papers of the British School at Rome* 33, 70-112.

Pala, C., 1976 *Nomentum. Forma Italiae. Regio I.* Rome.

Patterson, H., Di Giuseppe, H. and Witcher, R.E., 2004 'Three South Etrurian “Crises”: first results of the Tiber Valley Project', *Papers of the British School at Rome* **72**, 1-36.

Potter, T.W., 1979 *The Changing Landscape of South Etruria.* London.

Rajala, U., Harrison, A. and Stoddart, S.K.F., 1999 ‘The enhancement of the South Etruria Survey: GIS in the study of the research history of the Southern Faliscan area’, in L. Dingall, S. Exon, V. Gaffney, S. Laflin and M. van Leusen (eds), *Archaeology in the Age of the Internet. CAA97. Computer Applications and Quantitative Methods in Archaeology.* Oxford. CD-ROM.

Sharon, I., Dagan, Y. and Tzionit, G., 2004 ‘The [awful?] truth about GIS and archaeology’, in M. Iacovou (ed.), *Archaeological Field Survey in Cyprus: Past History, Future Potentials.* London, 151-62.

Terrenato, N., 2004 ‘Sample size matters! The paradox of global trends and local surveys’, in S.E. Alcock and J. Cherry (eds), *Side-by-Side Survey. Comparative Regional Studies in the Mediterranean World.* Oxford, 36-48.

van Hove, D., 2004 'Time and experience: taskscapes within GIS', *Internet Archaeology* **16**. http://intarch.ac.uk/journal/issue16/vanhove_index.html. Accessed 6 September 2007.

Ward-Perkins, J.B., 1961 'Veii. The historical topography of the ancient city', *Papers of the British School at Rome* **29**, 1-122.

Wise, A. and Miller, P., 1997 'Why metadata matters in archaeology', *Internet Archaeology* **2**. http://intarch.ac.uk/journal/issue2/wise_index.html. Accessed 6 September 2007.

Witcher, R.E., 1998 ‘Roman roads: phenomenological perspectives on roads in the landscape’, in C. Forcey, J. Hawthorne and R.E. Witcher (eds), *TRAC97. Proceedings of the Seventh Annual Theoretical Roman Archaeology Conference*. Oxford, 60-70.

Witcher, R.E., 2006 'Broken pots and meaningless dots? Surveying the rural landscapes of Roman Italy', *Papers of the British School at Rome* **74**, 39-72.

Witcher, R.E. and Craven, M., in press ‘“Much that has long been hidden…”. Reconstructing the survey methodology’, in H. Patterson (ed.), *Veii. The Historical Topography of the Ancient City. A Restudy of John Ward-Perkins’ Survey*. London.

List of figures

Figure 1: Location of the Biferno valley survey and *Forma Italiae* survey.

Figure 2: Larinum case study area. Spatial overlap of Biferno valley survey and *Forma Italiae* survey.

Figure 3: Distribution of Biferno valley survey findspots and soils in the Larinum case study area.

Figure 4: Graph of Biferno valley survey findspots by period in relation to soils in the Larinum case study area.

Figure 5: Distribution of Biferno valley survey villas in the Biferno valley and all other findspots.

Figure 6: Distribution of Biferno valley survey findspot 'dynamics' in the Larinum case study area. New, continuing and abandoned findspots at the Iron Age/Samnite period transition.

Figure 7: Distribution of Biferno valley survey findspots by number of occupations in the Larinum case study area.

Figure 8: Graph of Biferno valley survey findspots by period in the Biferno valley.

Figure 9: Distribution of Biferno valley survey and *Forma Italiae* survey findspots and topography (DEM curvature) in the Larinum case study area

Figure 10: Movement along the main ridge from the plain to Larinum.

Figure 11: Location of the South Etruria Survey (Eretum) and the Nomentum (*Forma Italiae*) survey.

Figure 12: Chronological variability of Roman findspots (Eretum and Nomentum).

List of tables

Table 1: Example of a Biferno valley survey record (#A1) with expanded details of gazetteer codes

Table 2: Example of a *Forma Italiae* (Larino) survey record (#42) with translation

Table 3: Summary of soil types within Larinum case study area (area/percentage) and summary of sampling by the Biferno valley survey (area/percentage of total)

Table 4: Chi-squared test results for Biferno valley survey findspots and soil types in the Larinum case study area (df = 6; α = 0.01)

Table 5: Average number of occupations by findspot interpretation (Biferno valley survey)

Table 1

Record A1	*Details*
LOC: 20 B3 109045	Located on map sheet 20; square B3; 109mm from left/east margin of map and 45mm from top/north
ENV: alt 475, top 1, asp 0, geo 1, lu 1	475m above sea level, on Basin Floor, with no Aspect (i.e. flat), on Alluvial soils; land under plough at time of survey
ARC: 4a	Small, light density scatter of artefacts
INT: sp (7/10, 13/14)	Sporadic (off-site?) scatter of generic Classical period and post-medieval/recent activity
Finds: ccw3, qf1, tl	3 sherds of Classical coarseware; 1 quern fragment; unquantified amounts of tile

Table 2

42. Villa? In località Pezza Don Pietro, m. 120 a SO di una casa di campagna (q. 414), sia a destra che a sinistra della mulattiera diretta verso la Masseria Petrucci sono sparsi frammenti di tegole, di mattoni e ceramica comune di epoca romana. È probabile che l'area (m. 200 x 60), in considerazione delle particolare collocazione topografica e del tipo di materiale rivenuto, sia stata occupata da una villa.
42. Villa? In the vicinity of Pezza Don Pietro, 120m south-west of the farmhouse (at spotheight 414m), on both the left and right of the track towards Masseria Petrucci, are scattered pieces of tile, brick and coarsewares of the Roman period. Considering the particular topographical location and the type of material recovered, it is probable that the area (200×60m) was occupied by a villa.

Table 3

Soils	Total case study area		BVS sample area	
	Km^2	%	Km^2	% sample
Alluvial (1)	15.88	12.35	14.53	91.46
Alluvial (2)	11.64	9.05	2.65	22.80
Clays	14.38	11.18	3.72	25.88
Limestone	41.44	32.22	15.30	36.93
Mixed sand/clays (1)	14.11	10.97	12.29	87.11
Mixed sand/clays (2)	19.59	15.23	12.10	61.76
Mixed sand/clays (3)	11.38	8.85	2.45	21.52
Total	128.61	100.00	63.04	–

Table 4

Soils	*Biferno valley survey*		
	Obs	*Expected*	*Calc χ^2*
Alluvial (1) (Pleistocene)	71	41.96	20.119
Alluvial (2) (Recent)	2	7.65	4.174
Clays	7	10.74	1.302
Limestone	31	44.17	3.927
Mixed sand/clays (1)	29	35.48	1.184
Mixed Sand/clays (2)	40	34.93	0.735
Mixed sand/clays (3)	2	7.07	3.639
Total/mean	182	182	35.080

Table 5

Interpretation	*Number of findspots*	*Total number of occupations*	*Average number of occupations*
Sporadic	465	551	1.18
Sporadic/Domestic	8	14	1.75
Domestic	181	310	1.71
Farmstead	83	173	2.08
Villa	15	54	3.60
Villa/Village	5	22	4.40
Village	4	12	3.00
Town	3	13	4.33

Figure 1

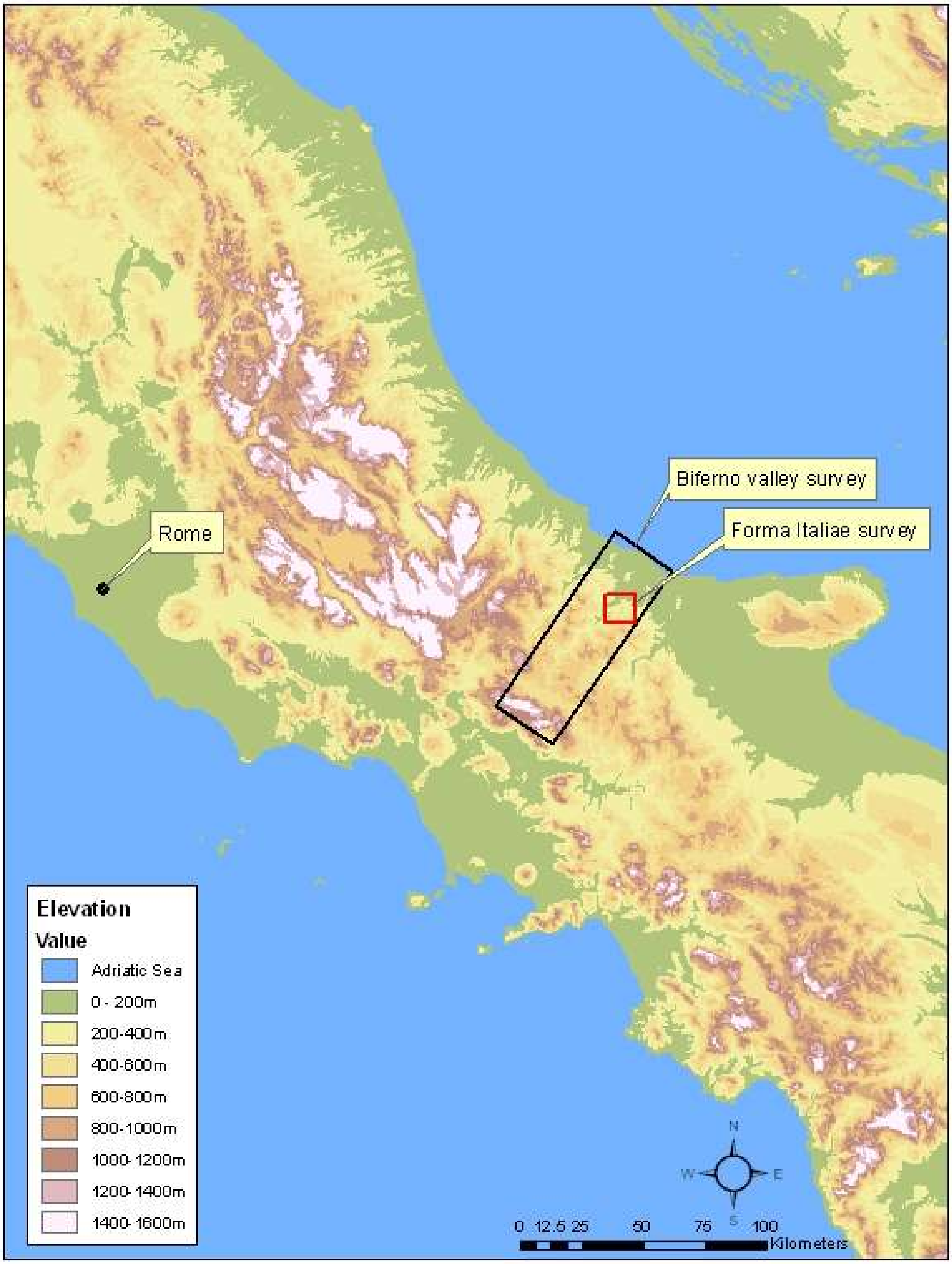

Biferno valley survey
Forma Italiae survey
Rome
Elevation
Value
Adriatic Sea
0 - 200m
200-400m
400-600m
600-800m
800-1000m
1000-1200m
1200-1400m
1400-1600m
N
W
E
S
0 12.5 25 50 75 100
Kilometers

Figure 2

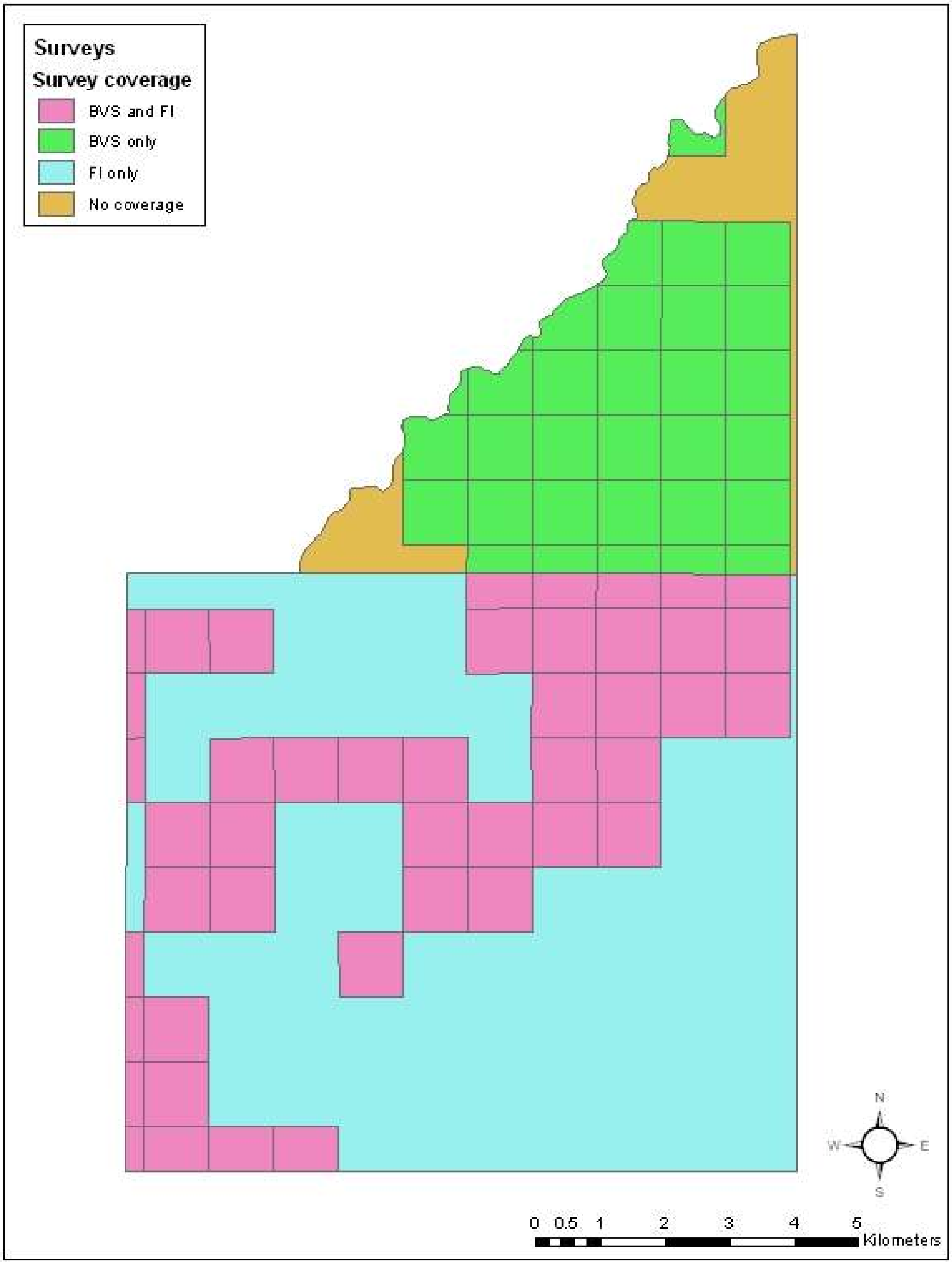
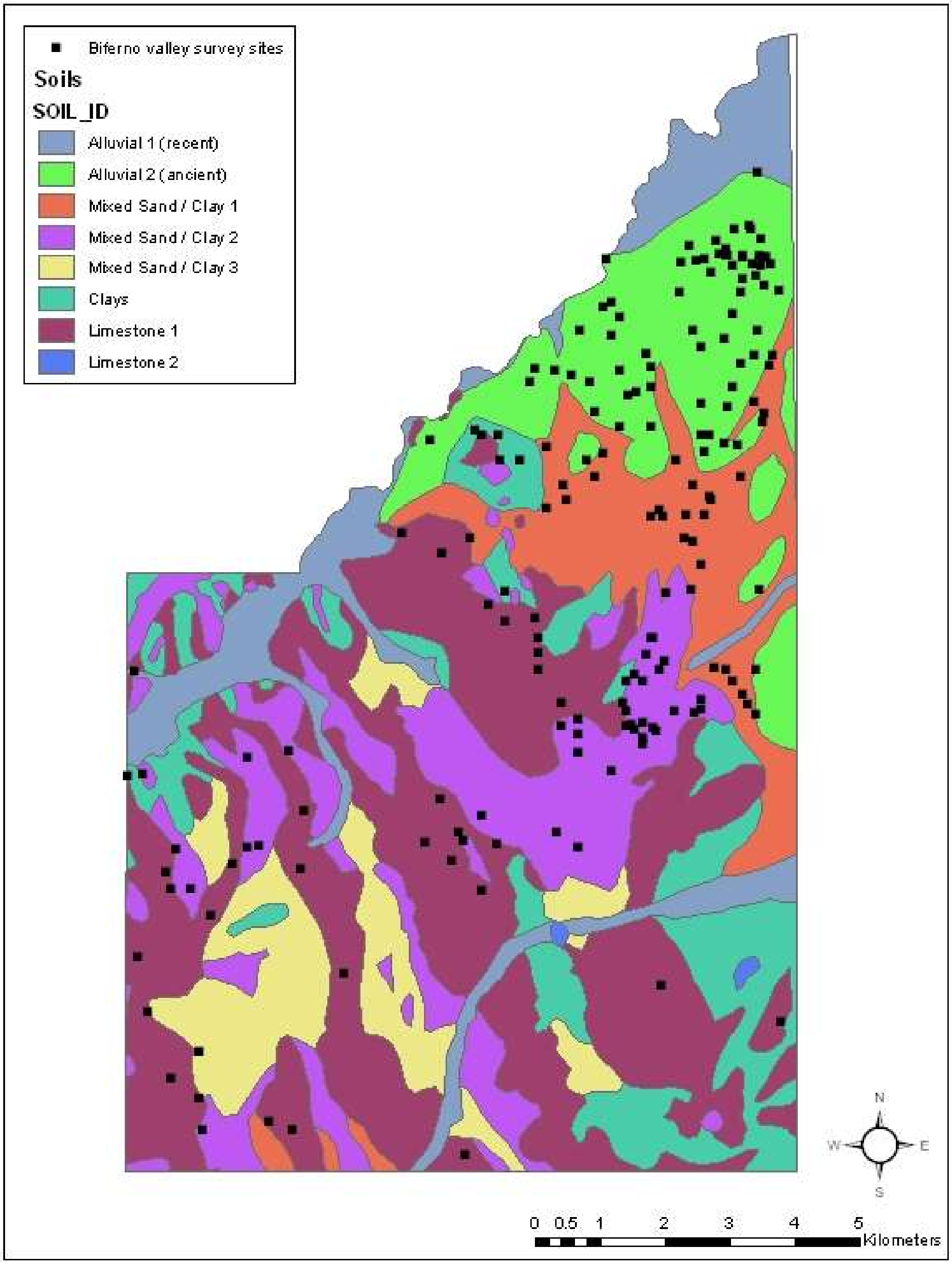

Surveys
Survey coverage
BVS and FI
BVS only
FI only
No coverage
N
W
E
S
0 0.5 1 2 3 4 5
Kilometers

Figure 3



Biferno valley survey sites
Soils
SOIL_ID
Alluvial 1 (recent)
Alluvial 2 (ancient)
Mixed Sand / Clay 1
Mixed Sand / Clay 2
Mixed Sand / Clay 3
Clays
Limestone 1
Limestone 2
N
W
E
S
0 0.5 1 2 3 4 5
Kilometers

Figure 4

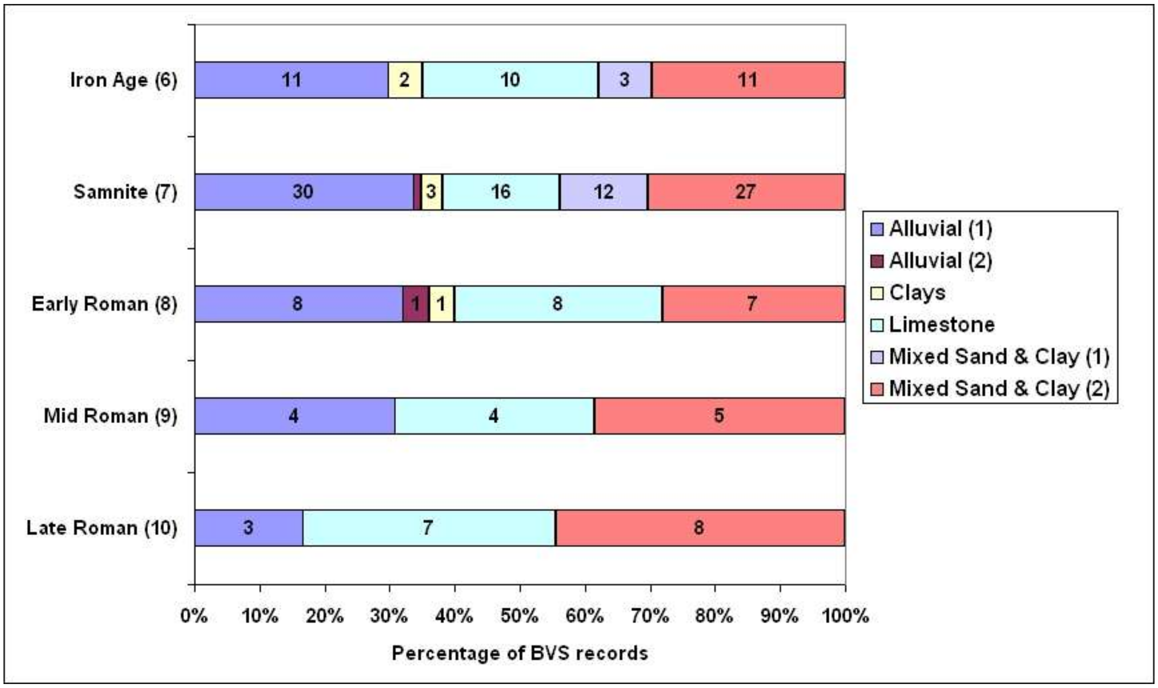

Iron Age (6)
Samnite (7)
Early Roman (8)
Mid Roman (9)
Late Roman (10)
11
2
10
3
11
30
3
16
12
27
8
1
1
8
7
4
4
5
3
7
8
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Percentage of BVS records
Alluvial (1)
Alluvial (2)
Clays
Limestone
Mixed Sand & Clay (1)
Mixed Sand & Clay (2)

Figure 5

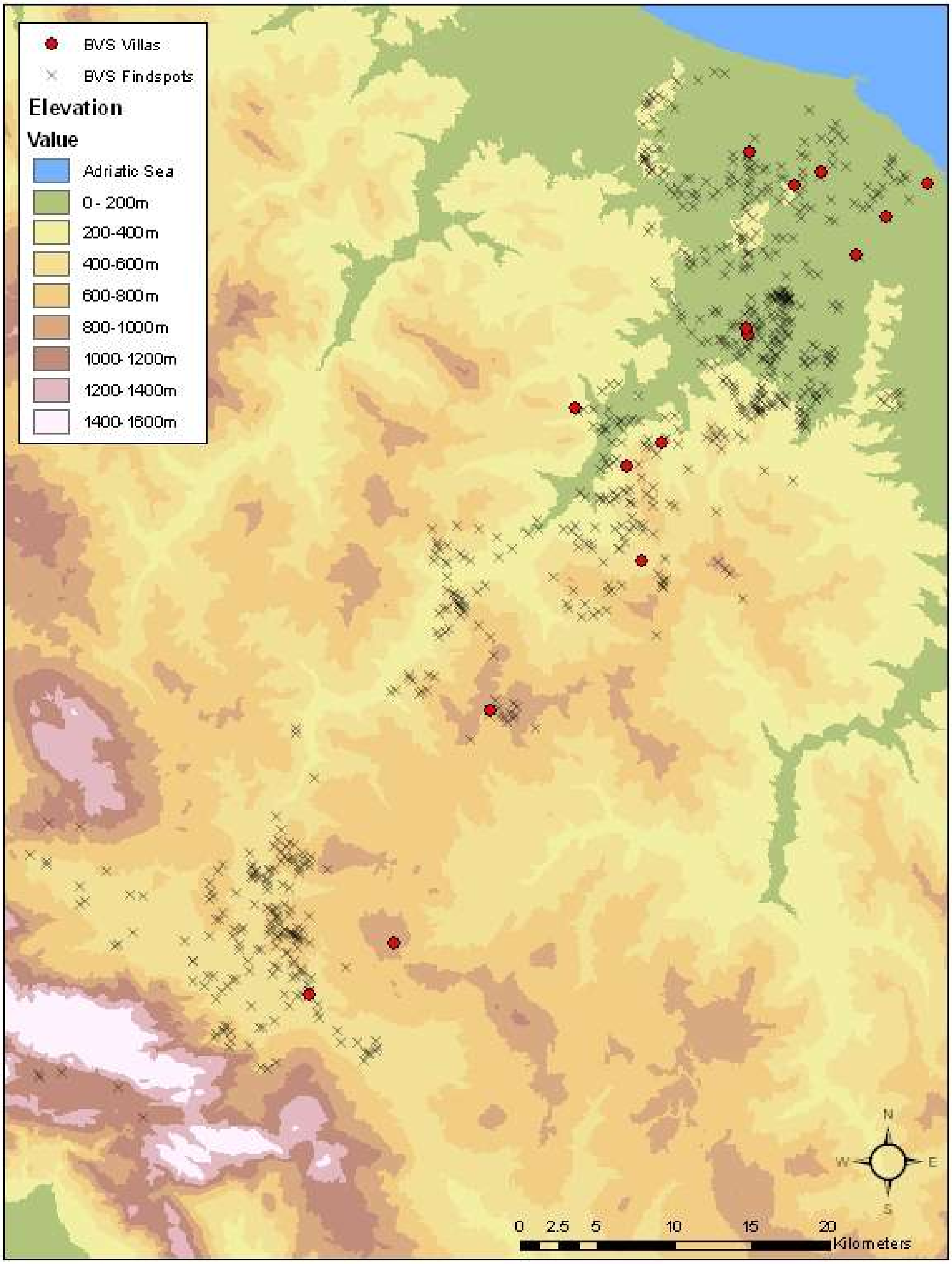
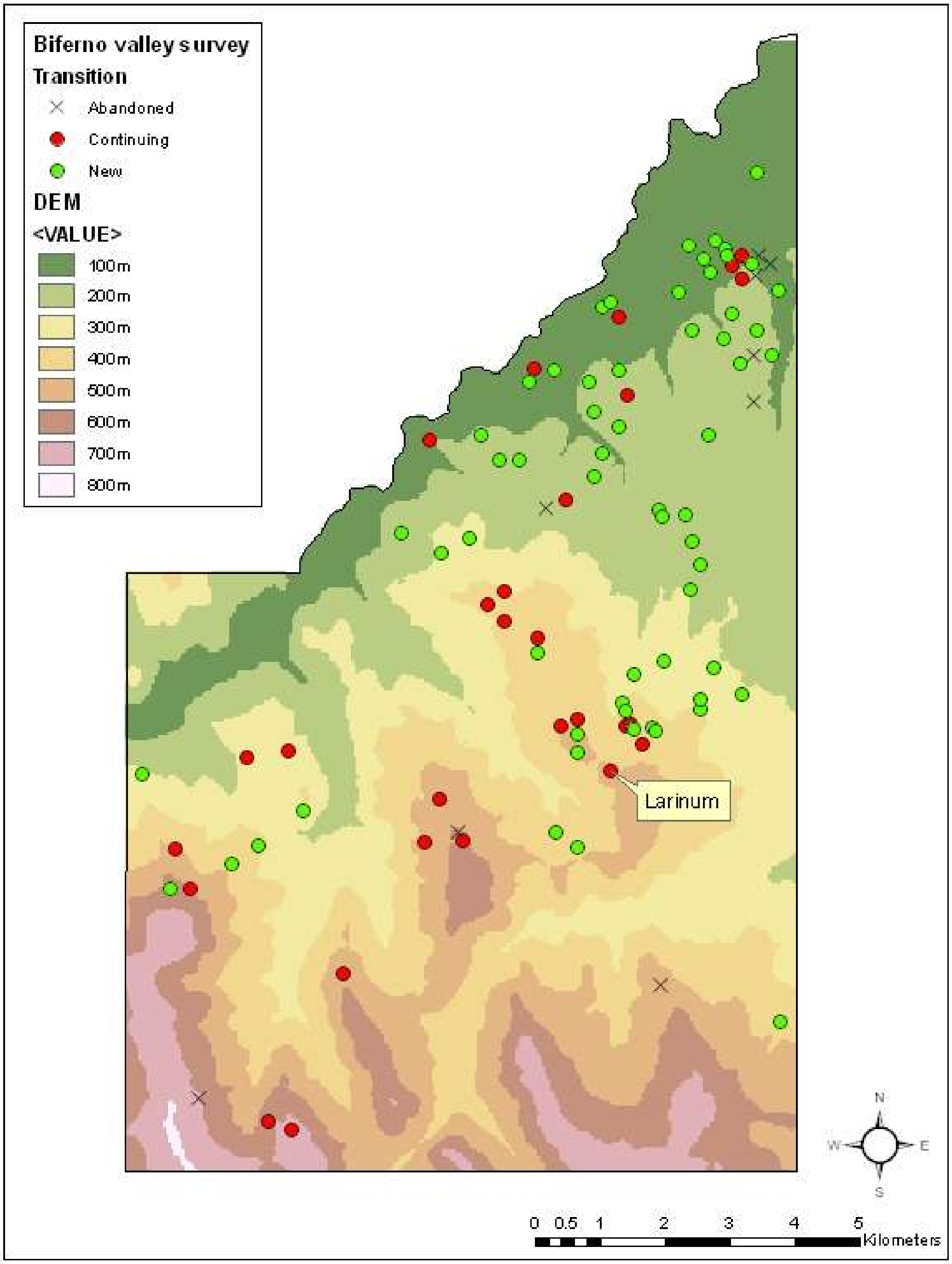
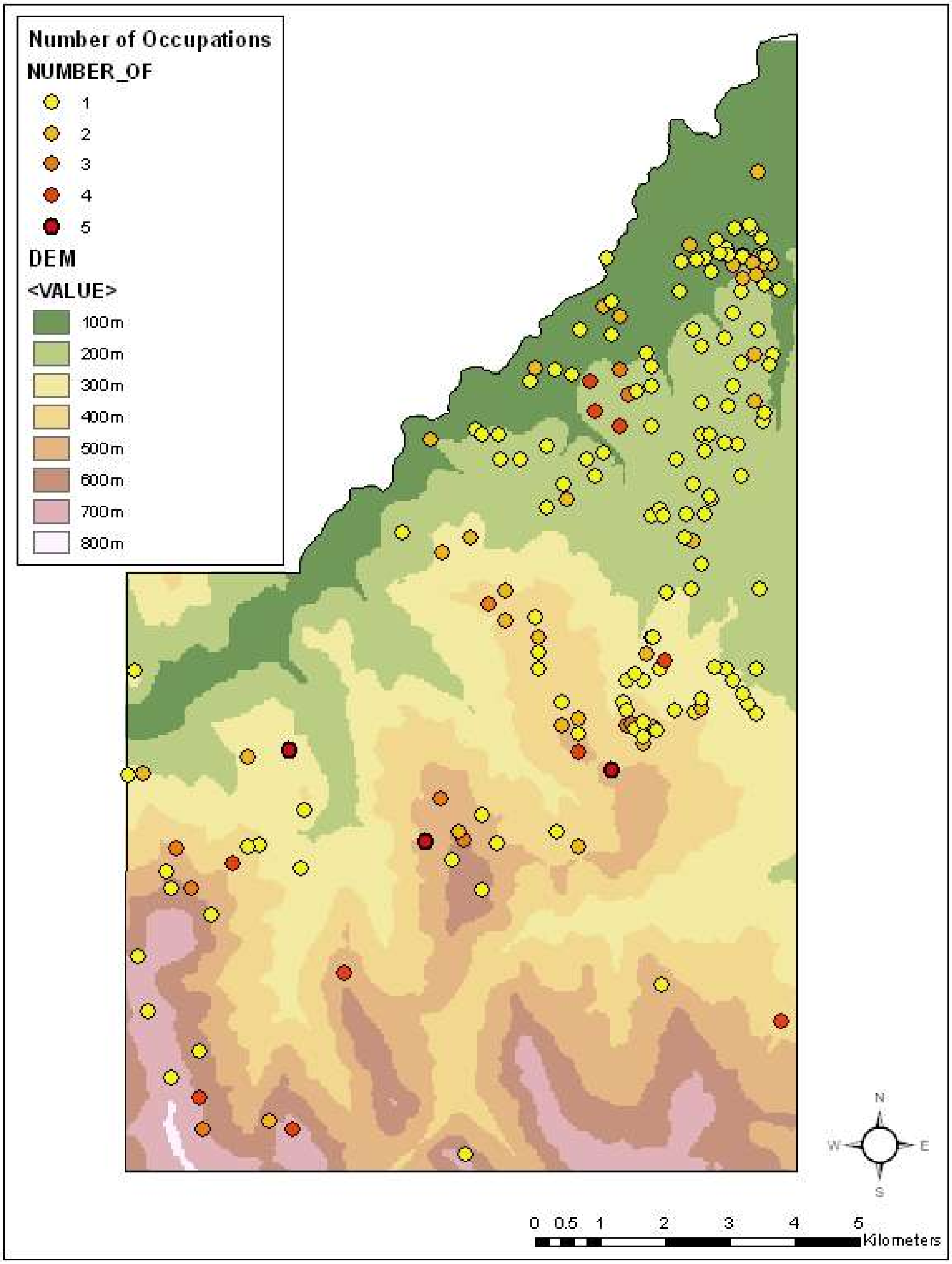

BVS Villas
BVS Findspots
Elevation
Value
Adriatic Sea
0 - 200m
200-400m
400-600m
600-800m
800-1000m
1000-1200m
1200-1400m
1400-1600m
N
W
E
S
0 2.5 5 10 15 20
Kilometers

Figure 6



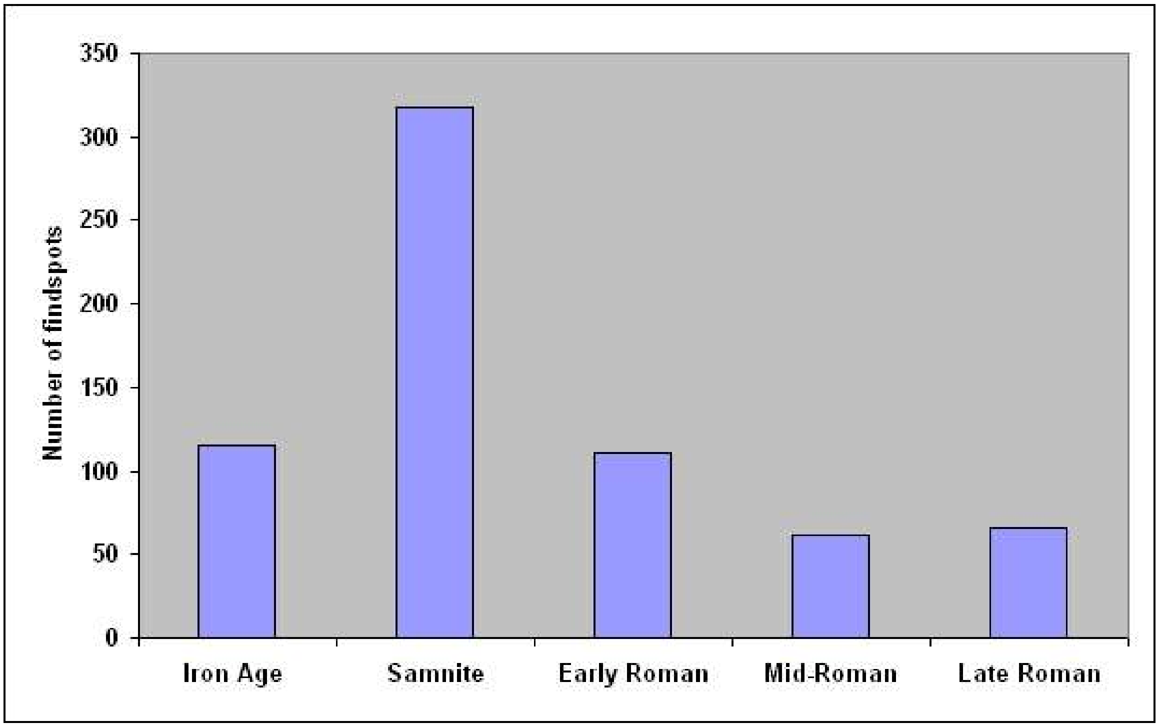
Biferno valley survey
Transition
Abandoned
Continuing
New
DEM
<VALUE>
100m
200m
300m
400m
500m
600m
700m
800m
Larinum
N
W
E
S
0 0.5 1 2 3 4 5
Kilometers

Figure 7



Number of Occupations
NUMBER_OF
1
2
3
4
5
DEM
<VALUE>
100m
200m
300m
400m
500m
600m
700m
800m
N
W
E
S
0 0.5 1 2 3 4 5
Kilometers

Figure 8



350
300
250
200
150
100
50
0
Number of findspots
Iron Age
Samnite
Early Roman
Mid-Roman
Late Roman

Figure 9

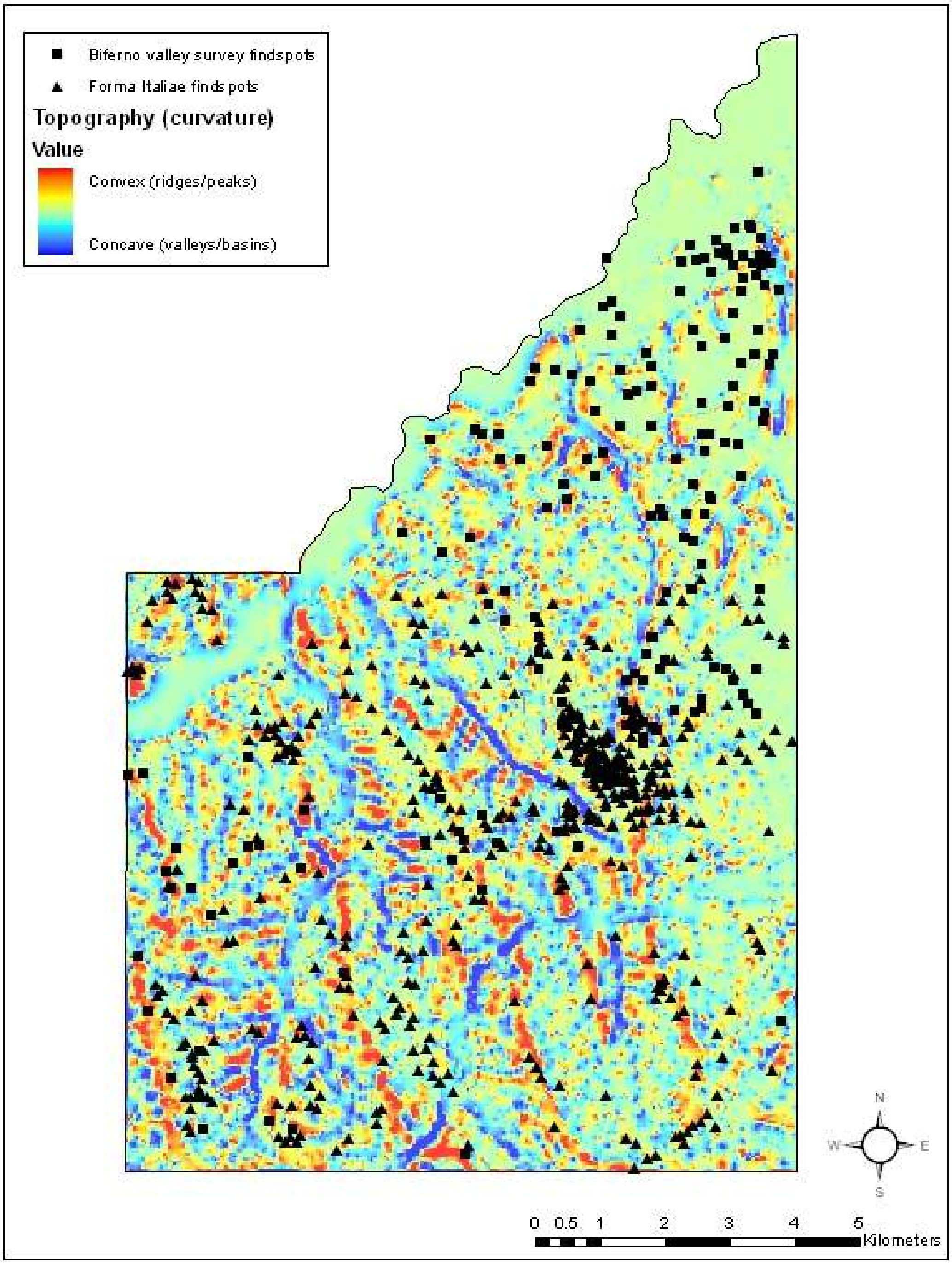

Biferno valley survey findspots
Forma Italiae findspots
Topography (curvature)
Value
Convex (ridges/peaks)
Concave (valleys/basins)
N
W
E
S
0 0.5 1 2 3 4 5
Kilometers

Figure 10

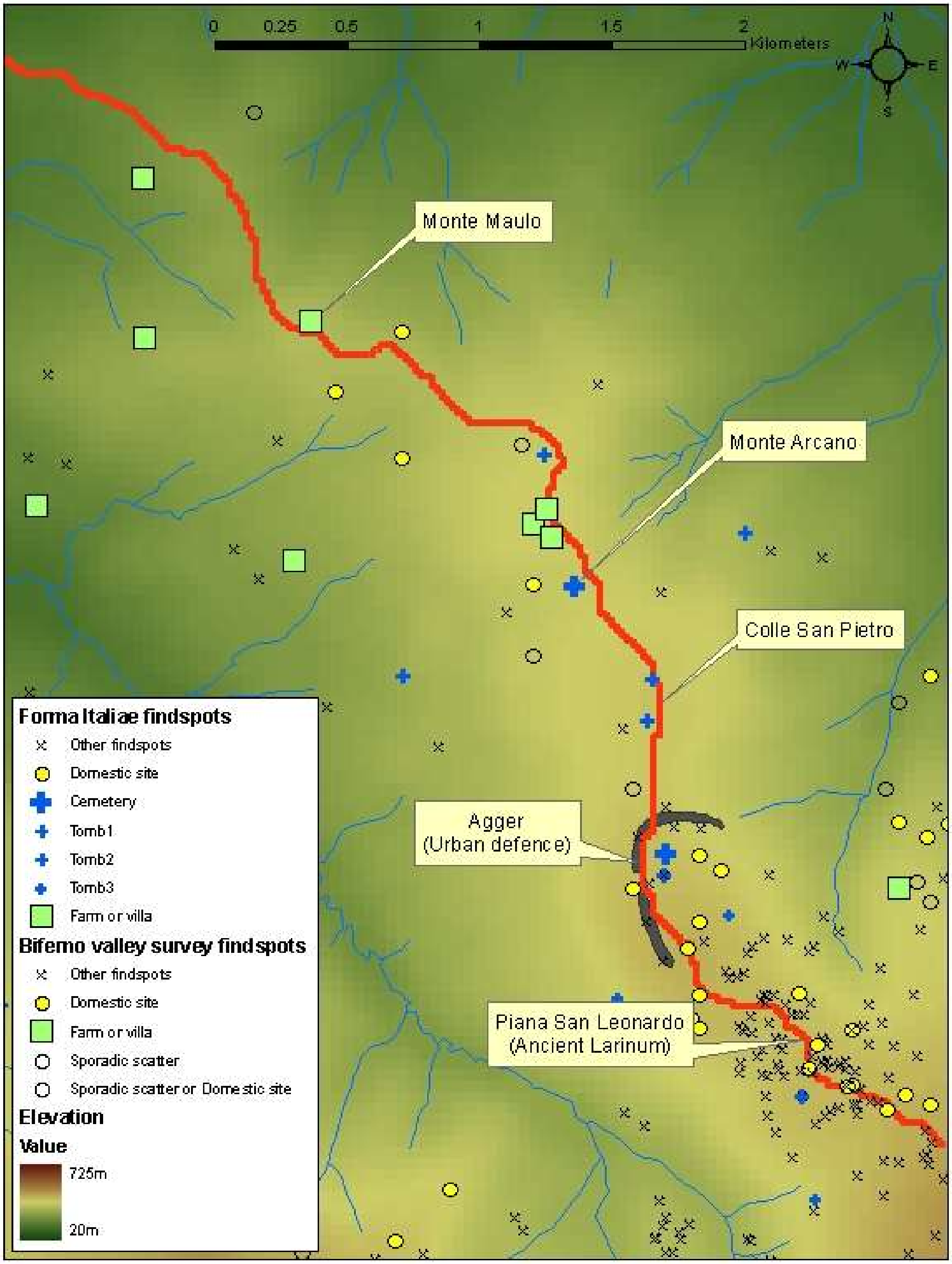

0
0.25
0.5
1
1.5
2
Kilometers
N
W
E
S
Monte Maulo
Monte Arcano
Colle San Pietro
Agger
(Urban defence)
Piana San Leonardo
(Ancient Larinum)
Forma Italiae findspots
Other findspots
Domestic site
Cemetery
Tomb1
Tomb2
Tomb3
Farm or villa
Bifemo valley survey findspots
Other findspots
Domestic site
Farm or villa
Sporadic scatter
Sporadic scatter or Domestic site
Elevation
Value
725m
20m

Figure 11

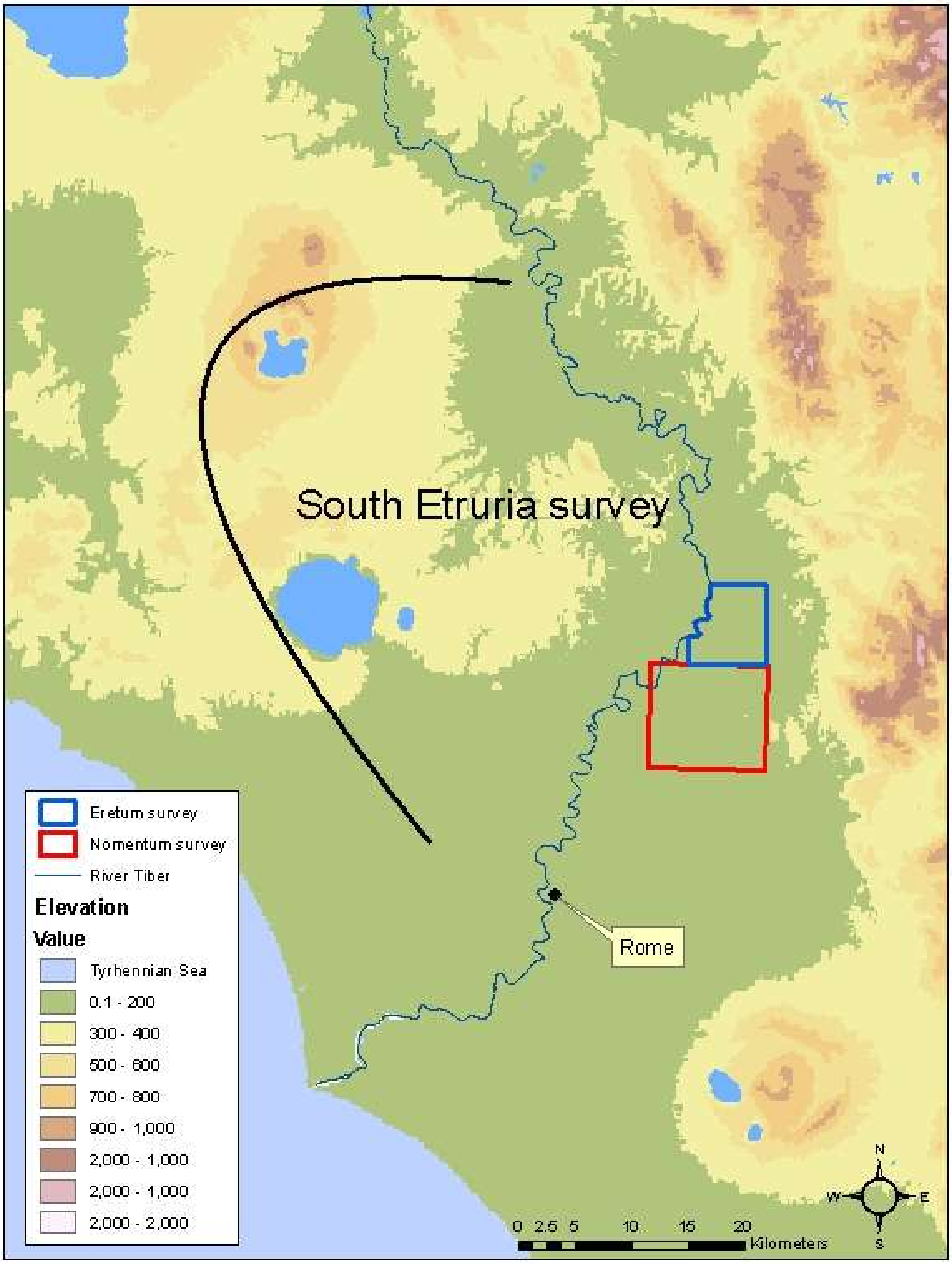
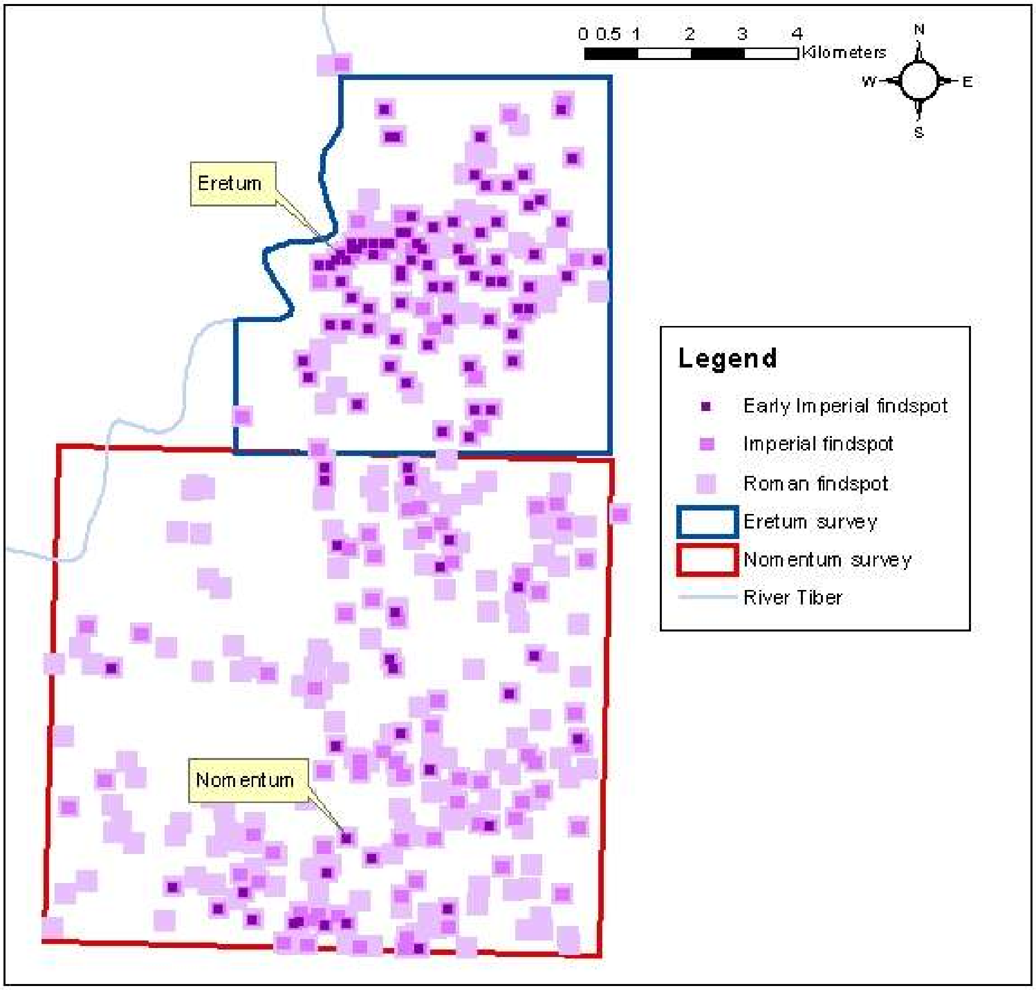

South Etruria survey
Rome
Eretum survey
Nomentum survey
River Tiber
Elevation
Value
Tyrhennian Sea
0.1 - 200
300 - 400
500 - 600
700 - 800
900 - 1,000
2,000 - 1,000
2,000 - 1,000
2,000 - 2,000
0 2.5 5 10 15 20
Kilometers
N
W
E
S

Figure 12



0 0.5 1 2 3 4
Kilometers
N
W
E
S
Eretum
Nomentum
Legend
Early Imperial findspot
Imperial findspot
Roman findspot
Eretum survey
Nomentum survey
River Tiber